Genome-wide identification of novel small RNAs in Pseudomonas aeruginosa

Gómez Lozano, María

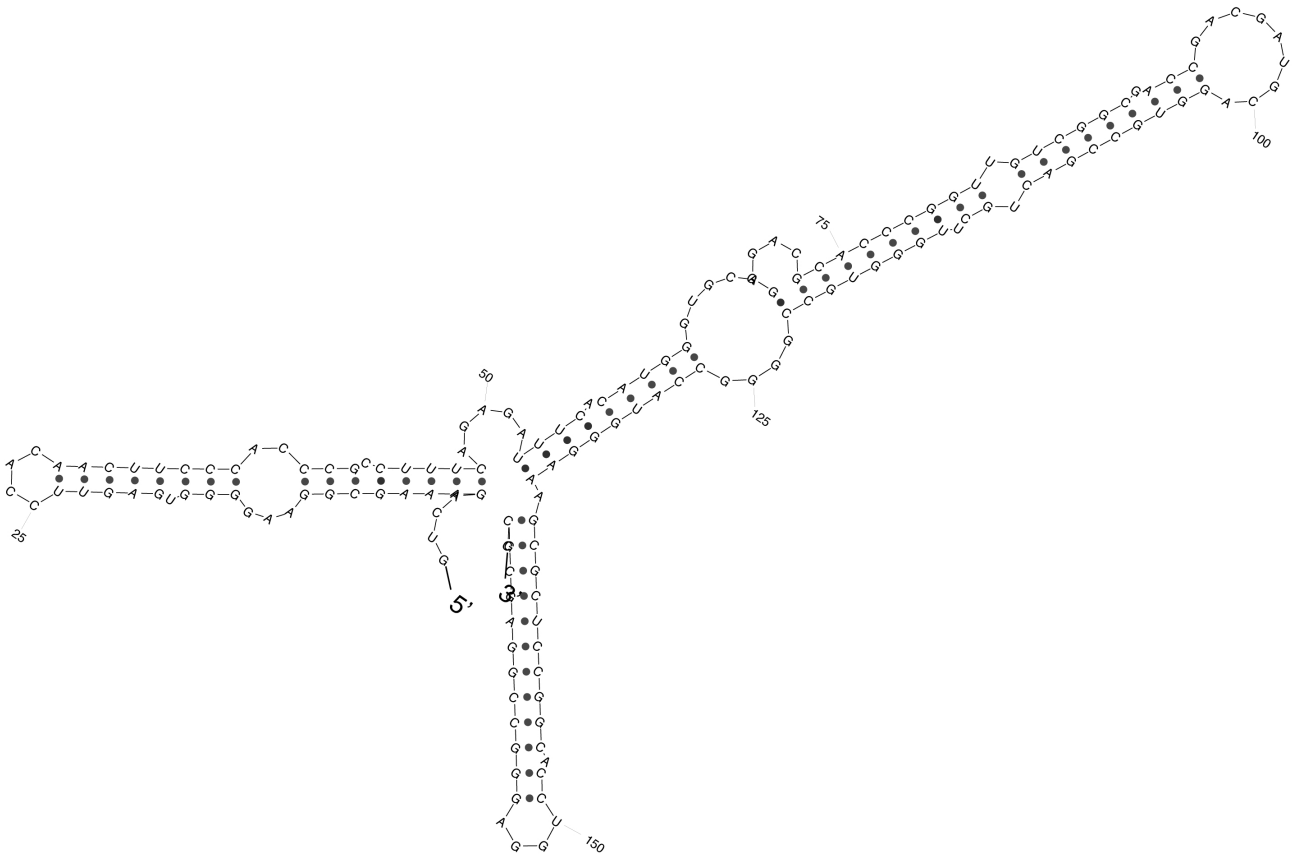
Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*

PhD thesis by

María Gómez-Lozano



Technical University of Denmark

Novo Nordisk Foundation Center for Biosustainability
Technical University of Denmark
Kogle Alle 6 2970 Hørsholm
Denmark

ISBN 978-87-91494-62-8

Cover illustration: Secondary structure of the primary transcript of OsiS RNA from *P. aeruginosa*

“A DNA sequence for the genome of bacteriophage φX174 of approximately 5,375 nucleotides has been determined using the rapid and simple 'plus and minus' method. The sequence identifies many of the features responsible for the production of the proteins of the nine known genes of the organism, including initiation and termination sites for the proteins and RNAs. Two pairs of genes are coded by the same region of DNA using different reading frames.”

Frederick Sanger, 1978

[Co-author]

Sanger, F., Coulson, A. R., Friedmann, T., Air, G. M., Barrell, B. G., Brown, N. L., et al. (1978). The nucleotide sequence of bacteriophage φX174. *Journal of Molecular Biology*, 125(2), 225–246. doi:10.1016/0022-2836(78)90346-7

PREFACE

This thesis is written as a partial fulfillment of the requirements to obtain a PhD degree at the Technical University of Denmark (DTU). The work presented in this thesis was carried out from October 2010 to September 2013. From October 2010 to December 2011 I worked at the Infection Microbiology Group (IMG), Department of Systems Biology at DTU, under the supervision of Professor Søren Molin and Katherine Long. From January 2012 to September 2013 I worked at the Novo Nordisk Foundation Center for Biosustainability (CFB) under the supervision of Professor Søren Molin. The work was financed by DTU, the Novo Nordisk Foundation and the Danish National Research Agency.

María Gómez-Lozano

Copenhagen, September 2013

ACKNOWLEDGEMENTS

Being a PhD student has been a great adventure with many joys and challenges. I would like to express my gratitude to all the people that have helped me in so many ways during these fantastic 3 years.

First of all, I must thank my supervisor Professor Søren Molin for giving me the opportunity to join his lab as a naïve MSc student. It is really hard to put into words how grateful I am for your excellent support and guidance, both scientifically and personally. As a scientist, you have taught me how to be rigorous, open, self-confident, creative, independent and passionate. Personally, you have always been tremendously encouraging and supportive. You are the kind of boss that listen to and respects other scientists, no matter whether they are renowned professors or undergraduate students. I am sure your guidance and personality will have a long lasting impact both on my career and on what kind of scientist I wish to become.

I would like to thank Katherine Long for her dedicated daily supervision during the first part of my PhD project. I will always be grateful for introducing me to the exciting topic of RNA regulation and for the time you spent helping me.

I would like to thank all the members of the Infection Microbiology Group (IMG), past and present: Lei Yang, Rasmus Marvig, Nicholas Jochumsen, Søren D Pedersen, Susse Hansen, Lars Jelsbak, Claus Sternberg, Anders Folkesson, Yang Liu, Liang Yang, Juliane C Thøgersen, Vinoth Wigneswaran, Martin W Nielsen, Lisse St. Clair-Norton, Sussane Koefoed, Sofia Feliziani, Trine Markussen, Fátima Yousef, Rasmus Bojsen, Alexandra Burleigh, Cristina Hierro, Adela Luján, Mariló Maciá, Eva Andresen and so many more. It has been a pleasure to work with all of you in the same group. I need to specially thank Lei Yang and Rasmus Marvig. Both have shown great enthusiasm and patience, Lei as my daily supervisor during my MSc thesis, and Rasmus as he taught me about bioinformatic analysis of sequencing data. I also need to thank Laurent Gautier and Khoa Do from the DmAc facility at DTU for their help and for allowing me the use of their labs during the first year of my PhD studies. During these three years, I also had the opportunity to supervise MSc

students Christian Bager and Monica Tulstrup. Thank you for your work and for teaching me how to be a teacher.

During my stay at the Novo Nordisk Foundation Center for Biosustainability (CFB) I have being very lucky to work with Martin H Rau, Carlotta Ronda, Anna Koza, José M Seoane, Mette Munk, Paula Tribelli, Alex T Nielsen, Christian B Jendresen, Helle K Johansen, Hemanshu Mundhada, Holger Døssing, Jerome Maury, Jiangfeng Zhu, Julia Hinderberger, Lasse H Lauridsen, Lone Riisberg, Markus Herrgard, Mikkel Lindegaard, Morten Nørholm, Morten Sommer, Morten T Nielsen, Neda F Parapari, Pernille Smith, Sailesh Malla, Sheila I Jensen, Simo A Jacobsen, Solvej Siedler, Susanna Seppala, Susanne R Pedersen, Ariane Zutz, Annika N Wallin, Isotta D'Arrigo, Jochen Förster, Margit Pedersen, Patricia Calero, Carlos Molina, Juan L Ramos, Estrella Duque, Margit Pedersen, and many many more. Mette Munk has been the perfect office mate, counselor and friend: funny, supportive, helpful, understanding, positive… You have always cared and been there for me, thanks! I need to thank Martin Holm Rau and Carlota Ronda for the fantastic scientific discussions and the funny times when going together to conferences. Anna Koza and José M Seoane, thanks for your friendship, for fantastic scientific discussions and great coffee breaks. I thank Morten Sommer, Alex T Nielsen and Markus Herrgard for his great suggestions about my work. Thanks also to Paula Tribelli for introducing me to *P. extremaaustralis* and for her help in the lab. Juan L Ramos and Estrella Duque are thanked for their help and great advice during their stay in our labs.

Last but not least, I want to thank my family and friends for their love and support: Mamá, Papá, Yayay, Tate, Patricia, Pati, Marta, Laura, Anneke, Achille, Stefano, Francesca, Raúl, Alberto, Juan, Alessandro, Eduardo, Tamara, Marcello, Jacopo. I need to thank my parents for supporting and loving me since… ever. Finally, I need to thank my husband, Rolando, for endless love, support and for making my life so colorful in so many ways.

ABSTRACT

Bacterial small regulatory RNAs (sRNAs) are known to have regulatory functions in a variety of processes including metabolic reactions, stress responses and pathogenesis in response to environmental signals. Recent genome-wide studies to identify sRNAs have been largely based on tiling arrays and RNA sequencing (RNA-seq) technologies. The latter approach, in particular, has revolutionized sRNA discovery by enabling interrogation of the transcriptome at unprecedented depths. The size and complexity of the *P. aeruginosa* genome suggests that it encodes many hitherto undetected sRNAs. In this study, RNA-seq is used to identify sRNAs in *P. aeruginosa* using a combination of three different sequencing libraries. Over 750 novel sRNAs (including intergenic and *cis*-encoded sRNAs) have been identified with this approach in this study. The results also reflect that although the use of three libraries increased the number of novel transcripts identified, there were significant differences in the subset of transcripts detected in each library, underscoring the importance of library preparation strategy and relative sRNA abundance for successful sRNA detection. These data will be useful for the study of regulatory sRNAs in *P. aeruginosa* and the approach described here may be applied to identify sRNAs in any bacterium under different growth and stress conditions.

In addition the role of sRNA OsiS was investigated. OsiS was identified in our genome-wide search of sRNAs in *P. aeruginosa*. OsiS is highly transcribed during oxidative stress conditions. We show that by inducing the expression of OsiS the levels of the sRNA PhrS are greatly reduced. PhrS activates the translation of the *pqsR* gene under low oxygen concentrations, which in turn activates the synthesis of the Pseudomonas quinolone signal (PQS). Thus, OsiS links the oxygen levels to the production of quorum sensing (QS) molecules. It is hypothesized that the interaction is by direct base-pairing between the two sRNAs, with a predicted recognition site of OsiS at the highly conserved-region of PhrS. However, more experiments are required to know the exact nature of the interaction between these two sRNAs. Notably, OsiS is, to the best of our knowledge, the first sRNA whose main function seems to be regulating the cellular levels of another sRNA.

RESUMÉ

Bakterielle små regulatoriske RNA (sRNA) molekyler er involveret i regulering i en række forskellige processer, heriblandt metaboliske, stress relaterede samt patogene som svar på eksterne signaler. Nylige teknologier anvendt til global genom-identificering af sRNA molekyler har overvejende været baseret på tiling arrays og RNA sekventering (RNA-seq). Specielt sidstnævnte teknologi har revolutioneret opdagelsen af sRNA molekyler ved at efterforske transkriptomet med et hidtil uhørt tilbundsgående detektionsniveau. Størrelsen og kompleksiteten af *P. aeruginosas* genom antyder at det kan kode for hidtil uopdagede sRNA molekyler. I dette studie har vi anvendt RNA-seq til at identificere sRNA molekyler i *P. aeruginosa* via en kombination af tre forskellige sekventeringsbiblioteker. Over 750 nye sRNA molekyler (inklusive intergeniske og cis-kodede sRNA molekyler) er blevet identificeret med denne metode. Studiet viser desuden at selvom brugen af tre forskellige biblioteker forøgede detektionen af antallet af nye transkripter, så var der betydelige forskelle i detekterede transkripter bibliotekerne imellem. Dette understreger at strategien for valg af bibliotek samt den relative mængde af sRNA molekyler har stor betydning for en succesfuld detektion af sRNA molekyler. Vi forventer at de præsenterede data vil være brugbare i studiet af regulatoriske sRNA molekyler i *P. aeruginosa* og at den beskrevne metode kan anvendes til identifikationen af sRNA molekyler i enhver given bakterie under forskellige vækst og stress forhold.

Et specifikt sRNA molekyle, OsiS, som blev identificeret i vores globale genom-undersøgelse af sRNA molekyler i *P. aeruginosa*, blev yderligere karakteriseret. Det viser sig at transkriptionen af OsiS er betydelig under oxidative stress forhold og vi viser at induktion af OsiS ekspression fører til stærkt reduceret ekspression af et andet sRNA molekyle, PhrS. PhrS aktiverer translationen af PqsR under lave ilt koncentrationer, hvilket aktiverer syntesen af Pseudomonas quinolone signal (PQS). OsiS repræsenterer derfor et link mellem oxygen niveauet og produktionen af quorum sensing (QS) molekyler. Det er vores hypotese at interaktionen mellem de to sRNA molekyler foregår via direkte base parring og OsiS har da også et forudsagt bindings-site i den meget konserverede del af PhrS. Yderligere eksperimenter er til gengæld nødvendige for at kunne bevise den eksakte interaktion mellem de to sRNA

molekyler. OsiS er desuden, ud fra vores viden, det første sRNA molekyle hvis overordnede funktion er at regulere det cellulære niveau af et andet sRNA molekyle.

PUBLICATIONS

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S. (2012). Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*. *Environmental Microbiology*, 14(8), 2006–2016. doi:10.1111/j.1462-2920.2012.02759.x

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S. (2013). Identification of bacterial small RNAs by RNA sequencing. *Methods in Pseudomonas aeruginosa. Accepted for publication.*

Gómez-Lozano, M., Marvig, R. L., Tulstrup, M. V. L., Tribelli, P., & Molin, S. (2013). Antisense small RNAs respond to osmotic, oxidative and antibiotic stress in *Pseudomonas aeruginosa*. *Manuscript in preparation.*

Gómez-Lozano, M. & Molin, S. (2013). Small RNA OsiS links oxidative stress to quorum sensing control in *Pseudomonas aeruginosa*. *Manuscript in preparation.*

CONTENTS

1. INTRODUCTION & OUTLINE

Historically, research on regulation of bacterial transcription focused on protein-based systems, and RNA was believed to be a relatively inactive information carrier. However, that has changed in the last decade due to the discovery of regulatory RNA activity. The regulatory role of certain bacterial RNA molecules has been known since the early 70s. However, until 2001, only about a dozen regulatory small RNAs (sRNAs) were identified in bacteria, representing biochemically abundant species and anecdotal discoveries. Technological advances such as DNA microarrays and RNA sequencing (RNA-seq) have recently allowed researchers to discover hundreds of sRNAs in all organisms studied so far. Bacterial sRNAs, which typically range from 70 to 500 nt in length, have been mainly found to function in post-transcriptional control of gene expression and control a variety of processes including metabolic reactions, stress responses and pathogenesis in response to environmental signals (reviewed in (1)). The main focus of this thesis was to identify novel transcripts in *Pseudomonas aeruginosa* and, in order to do so, the existing RNA-seq protocols have been optimized to increase the sensitivity of the method. As a result more than 700 novel sRNAs have been discovered in this clinically-relevant opportunistic pathogen. The optimized RNA-seq method can be used to study the transcriptome any bacterium in any growth condition.

This thesis is divided in six sections. **Section 2** (RNA sequencing and technology) introduces some general theory on transcriptomics and specifically on RNA-seq. The section serves as a fundamental theoretical background for the experimental investigations presented in the following chapters. **Section 3** (Bacterial small RNAs) provides basic information about the classification and the mechanisms of action of sRNAs in bacteria, and describes the traditional methods used to identify bacterial sRNAs and the current considerations when using RNA-seq to detect sRNAs. **Section 4** (*Pseudomonas aeruginosa*) focuses on the bacterium used for the investigations presented in this work, with an emphasis on antibiotic resistance, quorum sensing and known sRNAs. **Section 5** (Conclusions and perspectives) summarizes the main conclusions obtained from the research articles presented. Finally, **section 6** (Research articles) consists of the full-length research articles that were prepared as part of this PhD project.

2. TRANSCRIPTOMICS

2.1 RNA comes in more than 3 flavours

The “central dogma” of molecular biology was first enunciated by Francis Crick in 1958 (2), but was first sketched already in 1952 by James Watson (3) (Figure 1). The “central dogma” at that time postulated that the DNA information can be copied into RNA, and proteins can be synthesized using the information in the RNA in the cytoplasm. It was clear that although DNA was located in the eukaryotic nucleus, proteins were synthesized in the cytoplasm where there was no DNA, but there was lots of RNA (4). Most of the cytoplasmatic RNA could be found as discrete particles, which were later shown to be the ribosomes (5).

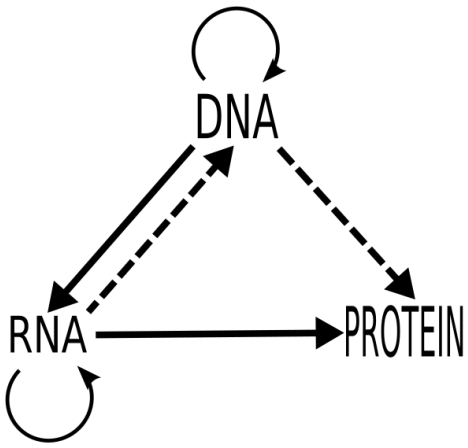


Figure 1. A diagram of the central dogma of molecular biology in 1958 (2). Solid arrows represent probable transfers, dotted arrows possible transfers.

Francis Crick predicted the existence of a second class of functional RNA, a molecule that acts as an adaptor between the triplet genetic code and the encoded amino acid (5). Interestingly, Crick hypothesized that the adaptor might be an RNA, because base pairing made RNA exceptionally fit for a role as a small, specific RNA recognition molecule (5). This hypothesis could be proven by Hoagland and colleagues, and Crick’s adaptors are now known as transfer RNAs (tRNAs) (6).

A few years later ribosomes were found to be composed largely of stable ribosomal RNAs (rRNAs), and to have no variety compared to the great variety found in genes

(7). It was shown that ribosomes were general-purpose RNA/protein machines, and that they were able to synthesize proteins when in contact with diverse unstable messenger RNAs (mRNA), which are only a small fraction of the total RNA population (7).

Therefore, in only a few years, RNA changed from coming in just one flavor (the messenger intermediate in the "central dogma") to having three flavors, all exclusively involved in protein synthesis: rRNA, tRNA, and everything else, which was thought to be mRNA. The non-ribosomal and non-transfer RNA fraction is complex, non-abundant and very unstable, and for decades there has been no interest in or no skills for investigating if it contained more than mRNA. Today, it is widely accepted that not all RNAs in the cell are involved in the synthesis of proteins as there are abundant small non-messenger RNAs (sRNAs), different from rRNA and tRNA, that regulate translation initiation, transcription termination, mRNA stability, or can sequester and modify proteins.

2.2 Transcriptomic technologies

Altering the levels of gene expression is one of the main mechanisms of cell regulation as it leads to all kinds of changes in the cellular machinery. The entire pool of transcribed sequences in a specific condition (termed transcriptome), provides enormous information about the physiological state of the cells. The first studies investigating transcription in cells used cDNA synthesis of specific mRNA transcripts, combined with cloning and Sanger sequencing (8). The transcriptomic field experienced a huge revolution when DNA microarrays were developed, providing a powerful tool to quantify expression of all annotated genes (9). DNA microarrays consist of immobilized gene-specific DNA probes that hybridize to their corresponding cDNA, producing a change in fluorescence. The measurement of fluorescence produced after hybridization correlates to the relative abundance of the transcript (Figure 2). The first cDNA microarrays had probes that hybridize to known or predicted mRNAs, and provided no information about transcription in intergenic regions or about transcription from the complementary strand of known mRNAs.

Several years later, the development of high-density tiled probes (tiling arrays) overcame this limitation. Tiling arrays contain probes that cover the entire genome of

an organism, so they are able to detect novel genes and sRNA species in intergenic regions in all organisms for which there was a tiling array available. Microarray-based studies provided the first global transcriptomic analyses in pathogens such as *Escherichia coli* (10), *Vibrio cholera* (11), *Borrelia burgdorferi* (12), *Chlamydia trachomatis* (13), *Chlamydia pneumoniae* (14), *Listeria monocytogenes* (15) and *Salmonella enterica* (16). In addition, transcriptomic studies using tiling arrays were able to provide accurate strand-specific information from 2006 thanks to the use of actinomycin D in the reverse transcription reaction to inhibit DNA polymerase activity (17).

While microarrays have been instrumental for the study of transcription, there are some major limitations to this approach. In microarray analysis, as with any other a probe-dependent method, the relative abundance of a given labeled transcript is inferred from the fluorescent signal produced following hybridization to immobilized probes. Transcript quantification is limited because the fluorescent intensity has both a lower (sensitivity) and upper (saturation) threshold. On the contrary, when using RNA-seq, which is a digital and probe-independent method, direct counting of sequencing reads for a given transcript has a much higher dynamic range. In addition, microarrays are usually designed to include sequences from a single strain, which will not be optimal for other strains of the same species as mismatches can significantly affect hybridization efficiency. There are microarrays that contain probes for multiple strains, but this may lead to a high background noise due to non-specific or cross-hybridization. Another limitation is the high cost of tiling arrays. Until recently, our interest was limited to mRNAs, tRNAs and rRNAs, and these RNAs could easily be probed with arrays at a reasonable cost. However, now it is known that mRNAs, tRNAs and rRNAs represent only part of the functional transcriptome. The study of the complete transcriptome at high resolution, including intergenic regions, requires microarrays with a very high density of probes (in the order of hundreds of millions), which currently exceeds the cost of RNA-seq experiments. Furthermore, the problem of cross-hybridization is increased in tiling arrays. Background noise due to cross-hybridization limits even more the dynamic range of microarrays (18). Also probe-based methods cannot discriminate between different mRNA isoforms or define transcript borders and splice junctions with high resolution.

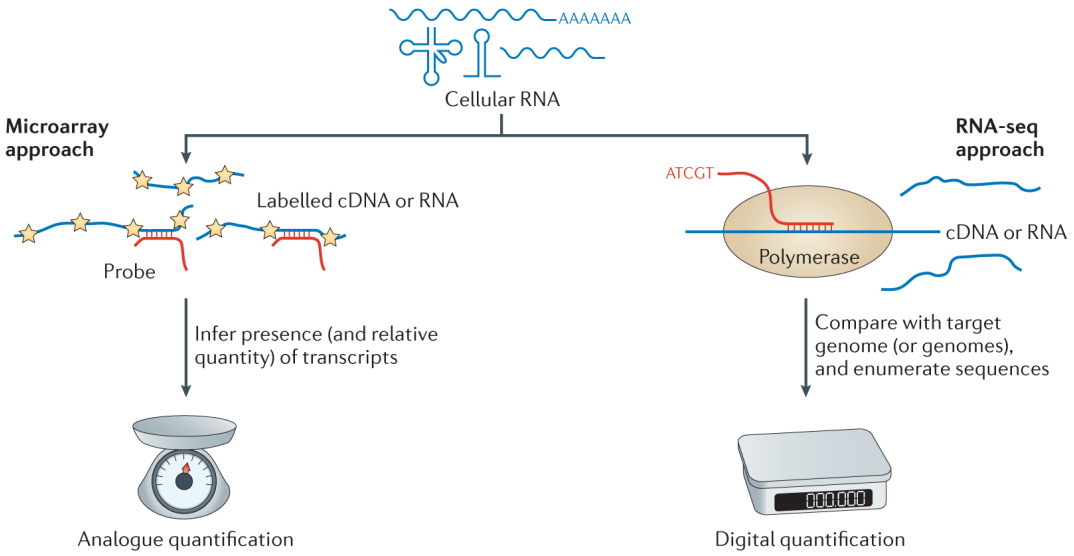


Figure 2. Fundamental differences between microarray and RNA-Seq technologies. Figure adapted from (19).

RNA-seq was introduced in 2008, and has already become the method of choice for studying transcriptomes as it overcomes many of the limitations of tiling arrays. It provides high resolution of the coordinates of transcripts, a higher dynamic range, lower cost, it does not have background noise due to cross-hybridization, and it is not a strain-specific method (reviewed in (20)). In addition to providing full-genome coverage, RNA-seq is very sensitive and has a high resolution. RNA-seq can identify novel transcripts and allows gene structure to be refined through accurate determination of transcript borders.

RNA-seq involves deep sequencing of cDNA generated from RNA preparations (reviewed in (21)). The design of library preparation methods is crucial when performing a high-quality RNA-seq experiment, as the data generated will depend greatly on the way libraries are prepared. The accuracy and precision of gene expression quantification and transcript identification are greatly influenced by cDNA library construction methods. The steps of a typical library preparation for RNA-seq are shown in figure 3.

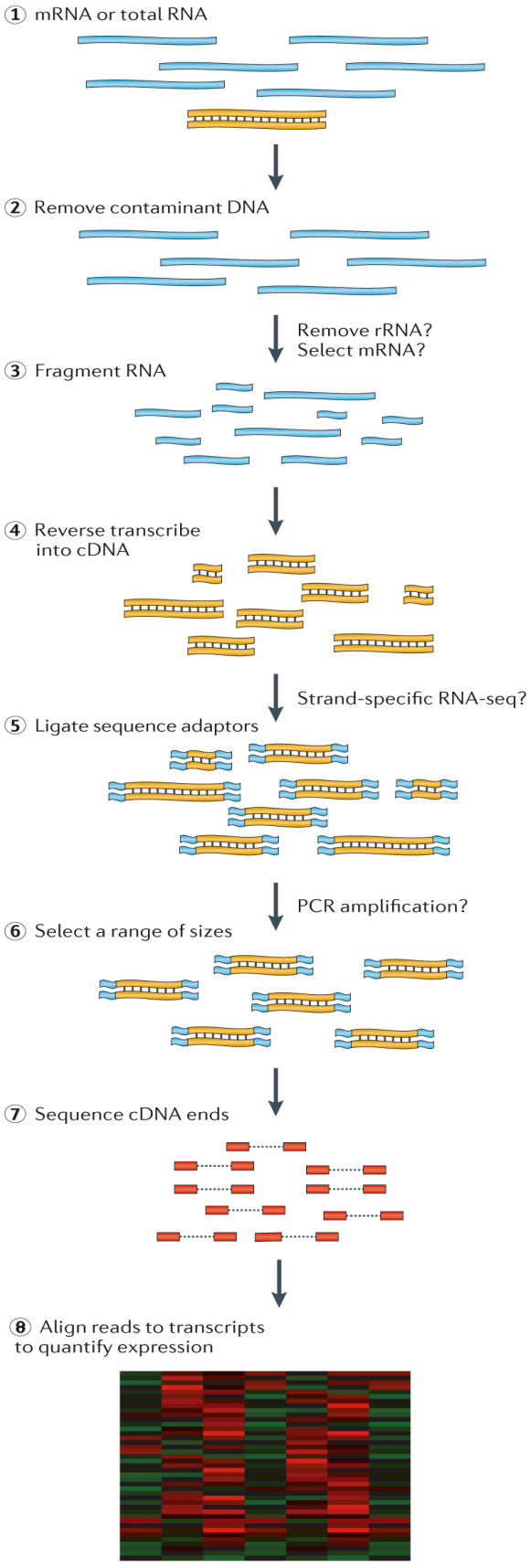


Figure 3. The data generation and analysis steps of a typical RNA-seq experiment. Figure adapted from (21).

Typically, total RNA is treated with DNase in order to remove any possible DNA contamination. In order to increase the genome coverage and the sensitivity of the experiment, rRNAs are usually removed during at this step of library preparation. Poly-A selection is very effective at enriching mRNAs in eukaryotes, but this approach does not retain sRNAs and mRNAs that lack a poly-A tail, and it is not suitable for prokaryotes. Instead rRNA can be removed by subtractive hybridization (Section 3.2 and Article 2 provide more details about successful rRNA-depletion in prokaryotes). If quantification of rRNAs is a goal of the study, then the sequencing of non-depleted libraries will be required. Following rRNA depletion, the RNA mixture is fragmented and converted into a library of cDNAs containing sequencing adaptors. Generally, the cDNA molecules are amplified before deep sequencing using primer complementary to the ligated adaptors. The cDNA library is then sequenced by high-throughput sequencers to produce millions of short reads from one end or both ends of the cDNA fragments. If both ends of the cDNAs are sequenced, then paired-end reads are generated. Paired-end reads are useful as they greatly reduce the complexity of transcriptome assembly. After the sequencing reaction has taken place, the obtained sequence reads are filtered by removing low-quality reads and artifacts, such as adaptor sequences and PCR duplicates. The filtered reads are then mapped onto a reference genome and assembled into transcripts. The expression level of each transcript is then estimated by counting the number of reads that align to each transcript. Section 3.2 provide information about optimal library preparation specific for sRNA identification, and Article 2 presented in this thesis describes in detail how to construct libraries specific for sRNA identification as well as how to analyze the data generated with them.

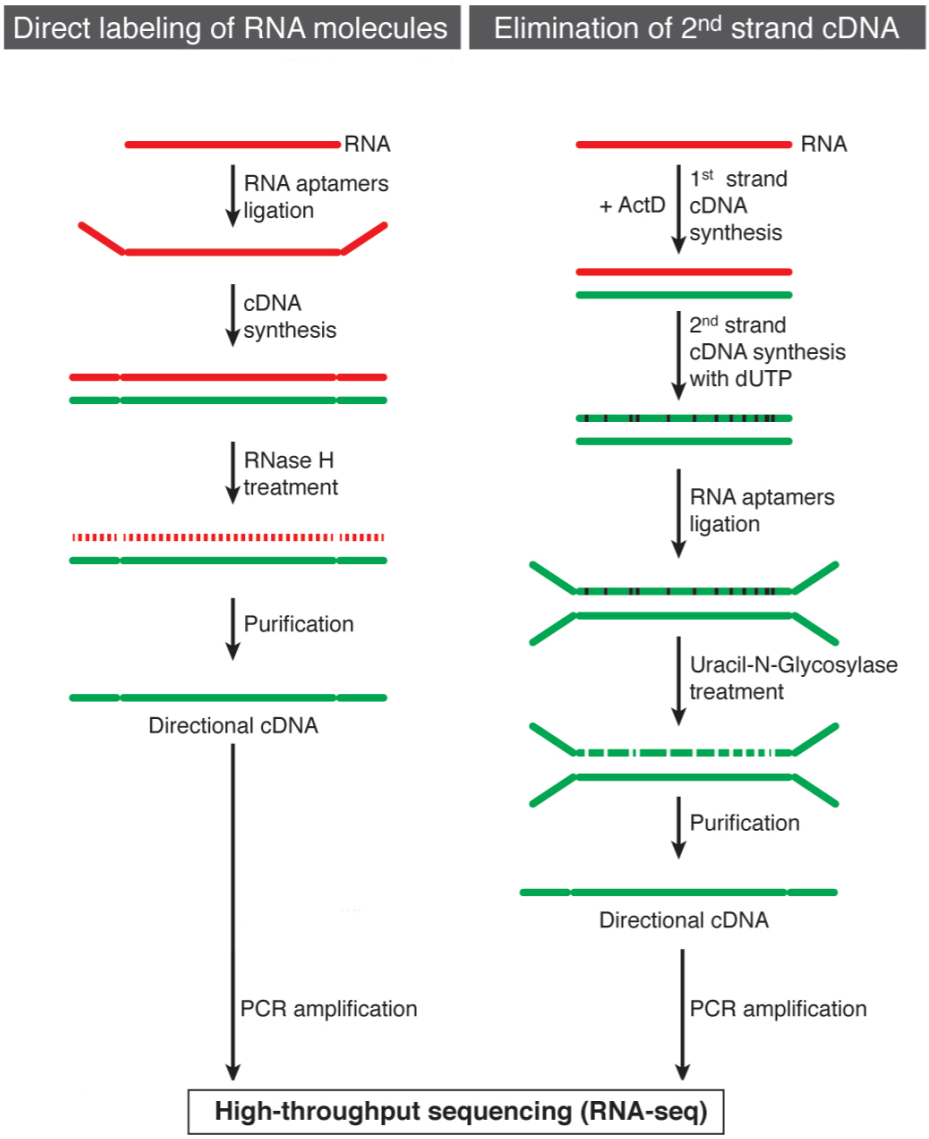


Figure 4. Flowchart illustrating technical improvements used to preserve the polarity of RNA molecules for transcriptome analysis. Figure adapted from (22).

An important step in the study of transcriptomes was the development of strand-specific RNA-seq protocols, which preserves information about the directionality of a transcript. Retaining strand-specificity is essential for studying for antisense transcription. In fact, by 2007, only about 30 antisense sRNAs (asRNAs) were known in bacteria (reviewed in (23)). After the introduction of strand-specific RNA-seq protocols the number of asRNAs detected has increased exponentially, with tens to hundreds of asRNAs detected in every bacterium studied (15,24-30). Strand-specificity is retained if the cDNA library is prepared using a directional protocol (Figure 4). During first-strand cDNA synthesis from RNA molecules by reverse transcription, unintended second-strand cDNA synthesis could occur using the first-

strand cDNA as a template. To avoid this, directional libraries can be prepared with the two following protocols: (i) first strand cDNA synthesis can be performed in the presence of actinomycin D, which specifically inhibits the DNA-dependent DNA polymerase activity of the reverse transcriptase enzyme (17). Then second strand cDNA synthesis is carried out in the presence of dUTP, which allows the selective removal of the strand with UNG (uracil-N-glycosylase) after ligation of 5' and 3' adaptors, leaving the first cDNA strand intact (31). Random oligonucleotides are used for the synthesis of first and second cDNA strands. (ii) Direct labeling of the RNA molecules with different adapters at the 5' and 3' ends before cDNA synthesis, which preserves the strand orientation of each RNA molecule (29,32). The transcripts are ligated to different RNA adapters in the 5’ and the 3’ ends using T4 RNA ligase. In this approach oligonucleotides complementary to the 5’ and 3’ adaptors are used for the synthesis of cDNA, considerably reducing the bias introduced by use of random primers to synthesize the cDNA.

3. BACTERIAL SMALL RNAs

The regulatory role of RNAs was first reported around 40 years ago in *E. coli,* when the plasmid- and transposon-encoded antisense sRNAs RNAI and CopA were found to negatively regulate plasmid copy number (33,34). However, there has not been much interest on this kind of regulation until the beginning of this century. Nowadays it is known that RNA-mediated regulation is found in all kingdoms of life, maybe in all species. In eukaryotes, microRNAs (miRNAs) are a large class of very short (21-25 nt) non-coding RNAs that affect translation and degradation of target mRNAs by antisense base pairing. The first miRNA was discovered in *C. elegans* in 1993 (35), but it was not until the early 2000s that miRNAs were recognized as a conserved class of biological regulators. Today miRNAs are known to be pivotal regulators of processes such as development and disease in eukaryotes (36,37).

In contrast to the eukaryotic miRNAs, bacterial sRNAs range in size approximately from 70 to 500 nucleotides (nt), though some sRNAs are shorter or longer than that. Bacterial sRNAs can modulate gene expression and most of them function through base-pairing interactions with mRNA targets that can affect translation initiation, transcription termination or mRNA stability (reviewed in (1,38)). There are examples of sRNAs involved in transcription reprogramming, carbon metabolism, iron homeostasis, cell envelope homeostasis and coordination of virulence (reviewed in (39)). At present, sRNAs have been found in all bacterial species investigated, and in some cases regulatory sRNAs may exceed protein regulators in number as well as diversity. Other RNA regulatory elements found in bacteria include riboswitches encoded in untranslated regions (UTRs) of mRNAs and CRISPR elements, and the reader is directed to fantastic existing reviews (40,41).

3.1 Mechanisms of action of bacterial small RNAs

Numerous bacterial sRNAs and all eukaryotic miRNAs regulate expression of mRNAs that are transcribed from a different genomic location, and are therefore named *trans*-encoded regulatory RNAs. There are also examples of *cis*-encoded sRNAs. They are transcribed from the reverse complementary strand of their target RNAs, thus generating fully complementary RNAs. Both *trans*- and *cis*-encoded sRNAs regulate the activity of a different RNA molecule, thus they are *trans*-acting regulators. There are also *cis*-encoded and *cis*-acting RNA elements, for example riboswitches encoded in UTRs of mRNAs.

A strand of DNA that is transcribed into an mRNA is called sense. Transcription from the opposite DNA strand (the antisense strand) of an mRNA-coding locus generates a *cis*-encoded antisense RNA that is fully complementary to the mRNA. However, *trans*-encoded antisense sRNAs also base pair to sense RNAs with limited complementarity. Thus, antisense sRNAs can be either *trans*- or *cis*-encoded.

In an effort to make the classification of sRNAs easier, *trans*-encoded and *trans*-acting sRNAs with limited complementarity with their targets are sometimes referred to as *intergenic sRNAs*, due to the fact that most of these sRNAs are encoded in intergenic regions. *Cis*-encoded and *trans*-acting sRNAs are normally referred to as *antisense sRNAs*, to emphasize that they are encoded from the antisense strand of the target RNA. These types of nomenclature, though not scrupulous, are easier and more widely used than the exact ones. In this thesis I sometimes refer to *trans*-encoded and *trans*-acting sRNAs as *intergenic sRNAs*, and to *cis*-encoded and *trans*-acting sRNAs as *antisense sRNAs* (asRNAs). I prefer not to use the term *non-coding RNAs* to refer to regulatory RNAs as some *trans*- and *cis*-encoded sRNAs also encode small peptides and are therefore dual-function sRNAs (reviewed in (42)).

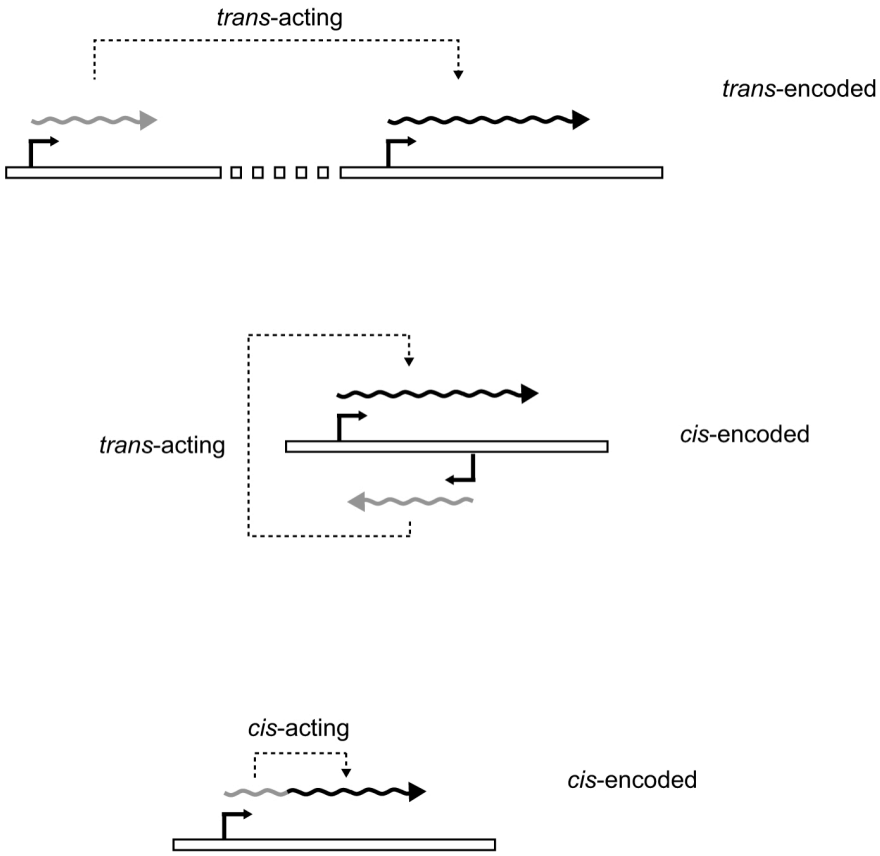


Figure 5. Regulatory RNAs can act on other RNA molecules (*trans*-acting) or on themselves (*cis*-acting), and can be expressed from a separate locus (*trans*-encoded) or from the same locus as the target RNA (*cis*-encoded). Regulatory RNA, mRNA, and DNA are shown in grey arrows, black arrows, and white boxes, respectively. Figure adapted from (43).

3.1.1 *Trans*-encoded sRNAs

After the discovery in the 1970s that the life cycle or copy number of bacterial transposons and plasmids is regulated by asRNAs, a similar regulatory RNA, MicF, was found to be encoded in the chromosome of *E. coli* (44). Unlike the asRNAs of transposons and plasmids, MicF was not transcribed from the DNA strand opposite its target gene *ompF*, which encodes the outer membrane protein F, but it was encoded in an intergenic region. Furthermore, MicF had incomplete and imperfect sequence complementarity to its target, but its base pairing strongly inhibited the translation of the *ompF* mRNA. Nowadays, most of the identified bacterial sRNAs are known to be encoded in different genomic loci (usually intergenic regions) than their target mRNAs and share only limited complementarity with their targets, like MicF.

Base pairing between the sRNA and its target mRNA can promote or inhibit

translation, and can decrease or increase mRNA stability. Most of the intergenic sRNAs characterized to date, bind to the 5' UTR of mRNAs most often inhibiting translation by base pairing with the ribosome-binding site (RBS) or upstream of the AUG codon of the mRNA (45,46). *Trans*-encoded sRNAs can also activate translation of their target mRNAs when base pairing of the sRNA disrupts an inhibitory secondary structure that sequesters the RBS (47-49). sRNA-mRNA duplexes are often degraded by RNase E or RNase III, the two major endoribonucleases in bacteria. The first, RNase III, specifically degrades double-stranded RNA complexes, which makes it an ideal candidate for processing asRNAs base paired with their targets. The asRNA-directed cleavage can generate two target RNA halves that are both less stable, both more stable, or differentially stabilized relative to each other. RNase E, which cleaves single-stranded RNA, is the other endonuclease involved in asRNA-directed processing. RNase E is a component of the degradosome, a multiprotein complex that affects mRNA stability and interacts with Hfq (reviewed in (50)). However it is not exactly clear how RNase E activity is stimulated by double-stranded RNA regions, though it is presumed that the 5' end of an base-pairing sRNA could provide a monophosphate, which has been shown to stimulate RNase E activity (51). *Trans*-encoded sRNAs can also increase the stability of mRNAs, by inducing cleavage by endonucleases that generates a mRNA with a different 5' structure that increases stability (52), or by protecting the 5' end of mRNAs from nucleases (53). For example, base pairing between the sRNA and its target mRNA might be able to block an RNase E recognition site, leading to increased stability of the target RNA. Theoretically, base pairing between a *trans*-encoded sRNA and its target could also promote transcription termination or antitermination, as has been found for some *cis*-encoded sRNAs.

There is little correlation between the genomic location of *trans*-encoded sRNAs and their target mRNA genes. In fact, several *trans*-encoded sRNAs are known to base pair with multiple mRNAs. This results from the fact that *trans*-encoded sRNAs have limited complementarity with their target mRNAs, usually in discontinuous patches, rather than extended regions of perfect complementarity, as for *cis*-encoded sRNAs. The region of base pairing between *trans*-encoded sRNAs and target mRNAs is typically 10–25 nt, but only some of the nucleotides seem to be critical for regulation. For example, although the SgrS sRNA can form 23 base pairs with the *ptsG* mRNA, only changes in 4 of these nucleotides in SgrS affects downregulation of *ptsG* (54).

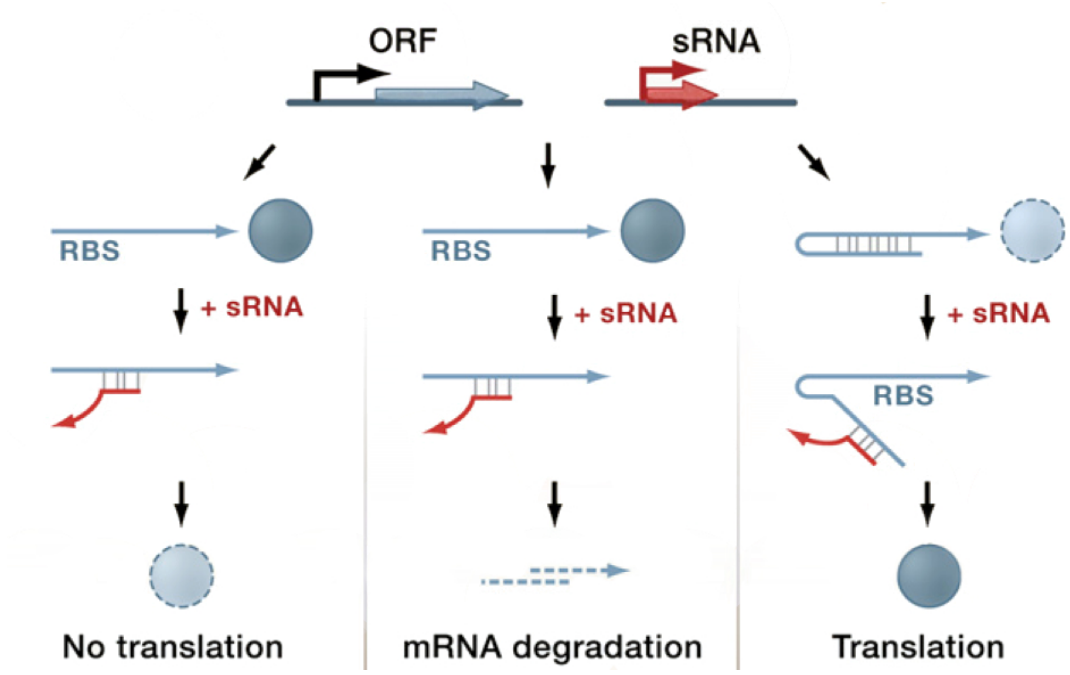


Figure 6. Mechanisms of action of *trans*-encoded sRNAs. Genes encoding *trans*-encoded sRNAs (red) are located separately from the genes encoding their target RNAs (blue) and only have limited complementarity. *Trans*-encoded sRNA can act negatively by base pairing with the 5'UTR and blocking ribosome binding (left panel) and/or targeting the sRNA-mRNA duplex for degradation by RNases (middle panel). *Trans*-encoded sRNA can also act positively by preventing the formation of an inhibitory structure, which sequesters the ribosome-binding site (RBS) (right panel). Figure adapted from (1).

In Gram-negative bacteria, the RNA binding protein Hfq is very frequently required for the function and stability of *trans*-encoded sRNAs (reviewed in (55)). However the detailed mechanisms by which Hfq promotes base pairing of sRNAs and their targets are still not well characterized. Hfq has been shown to facilitate base pairing by increasing annealing rates (56,57), by stabilizing sRNA-mRNA duplexes (58), and by promoting changes in secondary structure of one of the RNAs (59). Many *trans*-encoded sRNAs have been identified in studies that precipitate Hfq along with its bound RNAs (60-63). It is possible that there is a significant number of sRNAs that do not require Hfq to regulate mRNAs, but the most studied sRNAs so far are the ones that bind Hfq. Studies of Hfq in a range of organisms other than *E. coli* are necessary to uncover its full role as RNA chaperone. The *E. coli* ProQ protein has also been suggested to be an RNA chaperone, as it has a C-terminal domain with similar to Hfq (64). The YbeY protein, which is ubiquitous in bacteria, impacts gene expression similarly to Hfq in *Sinorhizobium meliloti* (65). Other proteins also may act as RNA chaperones bacterial species where there is no Hfq homolog.

3.1.2 *Cis*-encoded sRNAs

A significant number of protein-coding genes are also transcribed from the reverse complementary strand in a wide range of bacterial species (15,24-30). In general, overlapping transcription results in the generation of a *cis*-encoded asRNA whose size usually varies between various tens to hundreds of nucleotides (nt). By 2007, only about 30 bacterial asRNAs had been identified (reviewed in (23)). Antisense transcription detected in initial tiling array experiments was suspected to be an experimental artifact generated during cDNA synthesis. This problem was overcome by the development of strand- specific cDNA synthesis protocols (17). The number of reported asRNAs has lately exploded due to the use of RNA-seq. However characterization of the physiological function of individual asRNAs is growing at a much slower speed.

AsRNAs are known to affect the expression of the target gene by different mechanisms (Figure 7) (reviewed in (66)). (i) One is transcription interference, in which transcription from one promoter is suppressed by a second promoter located in the opposite strand. Transcription interference does not involve base-pairing and does not occur when the asRNA is provided in *trans*. This effect was first detected in two convergent bacteriophage promoters that produce transcripts which have an overlap of 62 bp at their 5' ends (67). It was observed that the stronger promoter significantly reduced the activity of the weaker promoter. When the promoters were oriented divergently, the transcription from both promoters was not affected despite the transcripts still maintained their regions of complementarity. Moreover, the introduction of a terminator before the convergent weak promoter resulted in reduced interference. These results led to the conclusion that the convergent orientation of the promoters was the source of the interference rather than base pairing. (ii) Antisense RNAs also can alter induce mRNA transcription attenuation, in which base pairing of the asRNA to the target mRNA causes premature transcription termination. In some cases base pairing of the asRNA to the mRNA has been shown to induce the formation of a terminator structure in the target mRNA. One example of this type of regulation is the asRNA RNAβ, encoded opposite to the *fatDCBA-angRT* iron transport-biosynthesis operon in the fish pathogen *Vibrio anguillarum* (99). Premature transcription termination results in increased levels of the *fatDCBA* fraction of the mRNA compared to the down-stream *angRT*, providing a mechanism for discoordinate expression within an operon. (iii) AsRNAs can also promote

changes in the target RNA stability by the same mechanism as intergenic sRNAs, where the asRNA either promotes or blocks degradation or cleavage of the mRNA by ribonucleases. As example, RNase III is known to be responsible for the of the plasmid-encoded *copT-copA* and *hok-sok* asRNA pairs (68,69). The AmgR RNA of *Salmonella enterica* induces degradation of the *mgtC* mRNA in a manner that requires RNase E but not RNase III (70). (iv) Just as *trans*-encoded sRNAs, asRNAs whose complementarity extends into the 5' UTR of their target RNA can impact on ribosome binding to the target mRNA (either positively or negatively) by affecting the target RNA structure.

In addition, regulating the expression of the opposite gene is not the only function of certain asRNAs. As intergenic sRNAs, some asRNAs encode small proteins (71), and some have the potential to act on multiple targets in *trans* (72-76). The number of existing asRNAs is far from complete, and other mechanisms of action will probably be found. Some possibilities are that antisense transcription serves to stabilize certain regions of the chromosome or as a defense against plasmids containing complementary regions.

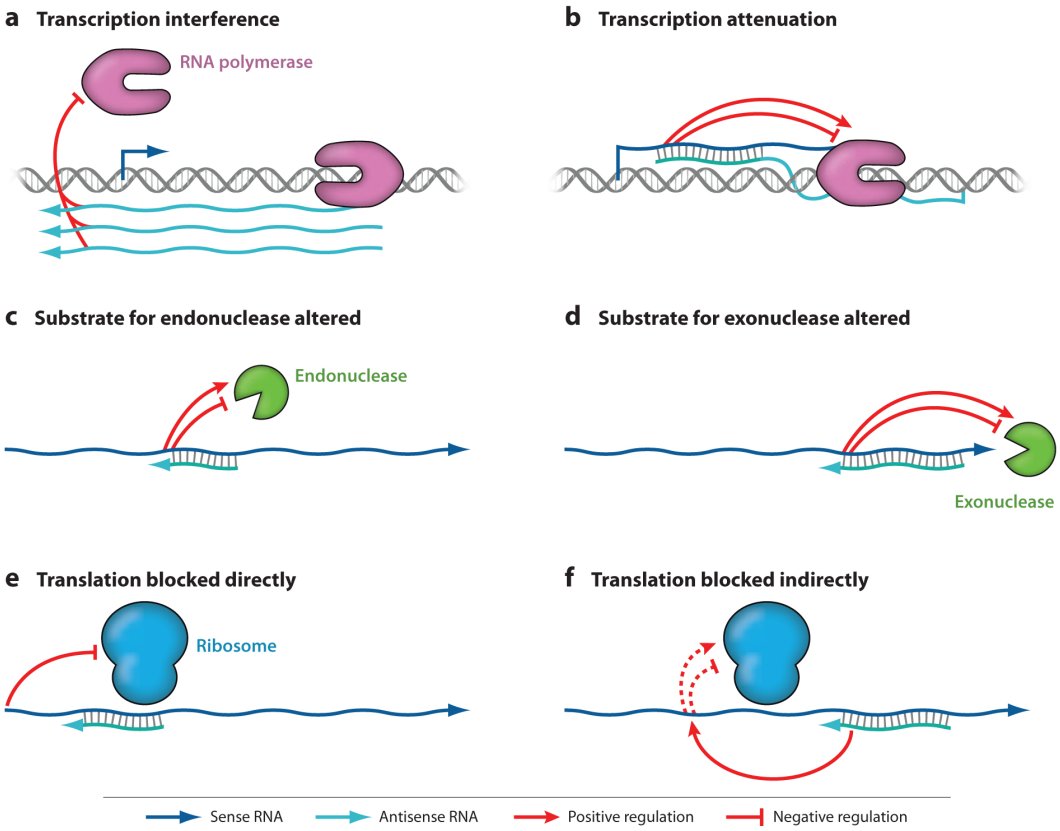


Figure 7. Mechanisms of action of antisense RNAs. Sense RNAs are depicted in dark blue while asRNA in light blue. Figure adapted from (66).

3.1.3 Protein-binding sRNAs

The number of sRNAs that directly regulate the activity of proteins is much lower than the number of intergenic sRNAs and asRNAs. However, their regulatory roles are by no means less powerful. sRNAs can regulate RNA binding proteins by containing the protein recognition sequence (molecular mimicry), often in multiple copies. This is the case of *E. coli* CsrB RNA, which has 18 binding sites for the RNA binding protein CsrA, which regulates mRNA translation and stability (reviewed in (77)). The *P. aeruginosa* CrcZ RNA is another example of this kind of regulation. CrcZ contains 5 CA-rich motifs which can bind up to 5 copies of the translation repressor protein Crc (78). Other sRNAs bind enzymes, and can inhibit, activate or modify protein activity. The most studied example is 6S RNA, which binds to the house-keeping form of RNA polymerase (σ^{70}-RNAP) by mimicking the secondary structure of DNA during transcription initiation (reviewed in (79)). The changes on transcription after binding 6S RNA are very complex, which indicates that σ^{70}-RNAP activity is modified rather than simply inhibited (80). It is expected that the number of sRNAs that bind to proteins will increase, and that more modes of regulation will be discovered. For example, it is proposed that sRNA binding to proteins might also modulate proteins by allosteric regulation or by tethering proteins close to each other (Figure 8).

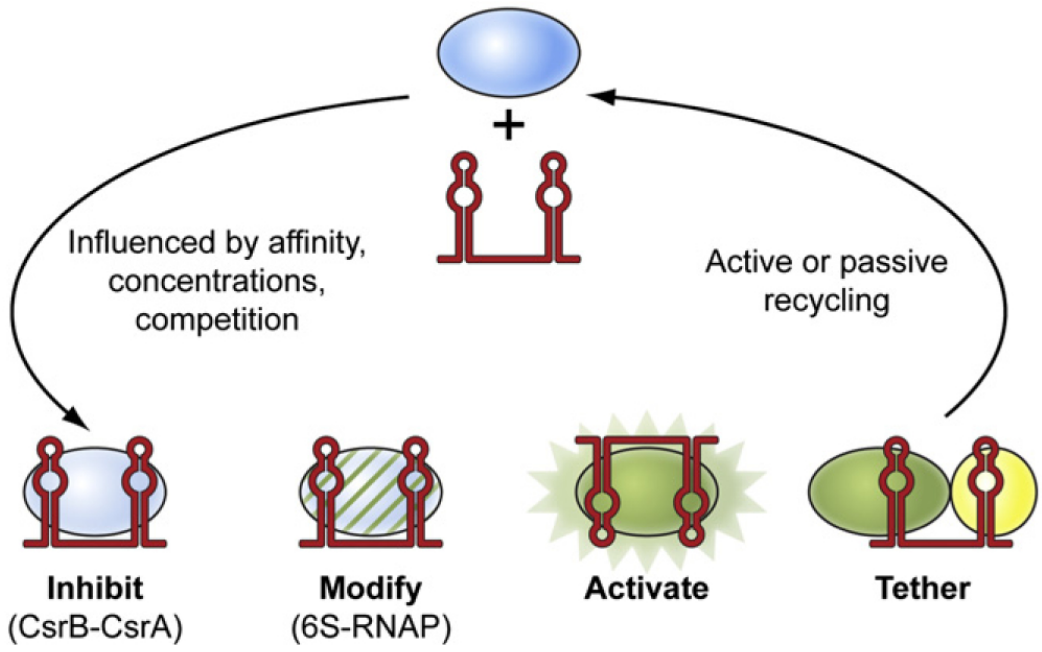


Figure 8. Mechanisms of action of sRNAs that modulate protein activity. Figure adapted from (81).

3.2 Identification of bacterial small RNAs

Until very recently, bacterial sRNAs have been overlooked mainly because there was no systematic technical skills to detect them. Despite the fact that the first bacterial sRNAs have been known for four decades, earlier studies to identify sRNAs have often depended on serendipity, direct detection due to high abundance, protein co-purification, and computational predictions based on sequence conservation (reviewed in (82)). Computational predictions of sRNAs often search exclusively within intergenic regions and only detect sRNAs that fall into a certain definition, based on a combination of sequence conservation, specific promoters and Rho-independent terminators (83-86). Although various sRNAs have been identified in bioinformatics predictions, sRNAs that are transcribed within or antisense to already annotated loci are missed. In addition, sRNAs whose 3' ends are formed by Rho-dependent terminators or processed by RNases are missed by computational methods that search for Rho-independent terminators. Similarly, studies that rely on chaperone-binding to find sRNAs will miss chaperone-independent sRNAs.

Until 2001 only a dozen of sRNAs were known in bacteria, 10 of them identified in *E. coli* (reviewed in (87)). Most of these sRNAs were identified serendipitously while analyzing flanking protein-coding genes. In 2001, the first genome-wide systematic searches for sRNAs were performed in *E. coli* by three independent groups, and more than 50 sRNAs were identified (62,88,89). These studies predicted the existence of sRNAs based on sequence conservation between *E. coli* intergenic regions and those of the closely related bacteria *Salmonella typhimurium* and *Yersinia pestis*, for which complete genome sequences had been published at the time. By 2007, only 140 bacterial sRNAs had been discovered (reviewed in (90)).

The first transcriptomic studies investigating sRNAs used cDNA synthesis from RNA samples, combined with cloning and Sanger sequencing (60,91). Recent genome-wide studies to identify sRNAs have been based on tiling arrays and, since 2008, primarily on RNA-seq. As mentioned in section 2.2, RNA-seq has many advantages over tiling arrays, and has enabled the study of the transcriptome at unprecedented depths and in a strand-specific manner. RNA-seq has revealed several hundreds of previously undetected transcripts in diverse bacteria including *Bacillus subtilis* (92), *Chlamydia trachomatis* (28), *Helicobacter pylori* (24), *Legionella pneumophila* (93), *Listeria monocytogenes* (94,95), *Salmonella enterica* (96-98), *Streptomyces*

coelicolor (99), *Synechocystis* (100), *Vibrio cholerae* (91,101), *Yersinia pseudotuberculosis* (102), *Escherichia coli* (103,104), *Mycobacterium tuberculosis* (105), *Staphylococcus aureus* (29), and *Agrobacterium tumefaciens* (106).

Special attention must be paid to the way RNA-seq experiments are designed, as different protocols will yield a very different number of detected sRNAs. Many of the RNA-seq studies published so far are not optimal for the identification of sRNAs, even if some are identified. One of the first considerations when designing an RNA-seq experiment to study sRNAs is the method used to extract total RNA. Historically RNA extraction methods have been optimized to exclude sRNAs under the assumption that they were products of degradation of no interest or importance. Many commercially available kits for RNA isolation do not quantitatively retain the short RNAs. This is the case for most of the kits that use spin columns. Spin-column kits provide very quick and efficient methods for extracting high-quality RNA, but are not suitable to study sRNAs since their size exclusion limit is too large. Transcriptomic studies that use this kind of RNA extraction cannot detect most of the sRNAs transcribed by both prokaryotic and eukaryotic cells. There are commercially available spin-column kits specially designed for the study of eukaryotic miRNAs (20-22 nt). These kits are not suitable for the study of prokaryotic sRNAs, as these typically range in size from 70 to 500 nt. Methods as the hot-phenol protocol or the guanidinium thiocyanate-phenol-chloroform extraction allow retention of all RNA species (see Article 2 presented in this thesis for details). Another important consideration is that strand-specific library preparations are necessary for the accurate determination of the direction of the transcripts, as well as the identification of antisense transcription events, as mentioned in section 2.

Other consideration when designing RNA-seq experiments is whether or not it is desirable to include highly abundant rRNAs in the library preparation. The relative abundance of mRNAs in bacterial total RNA can be as low as 5%, being the rest rRNAs and tRNAs (107). This results in a major technical challenge for RNA-seq. Unlike eukaryotic mRNAs, which can be selectively synthesized into cDNA by virtue of their poly-A tails (108), bacterial and archaeal cDNA synthesis results predominantly in rRNA sequences (109). To improve mRNA and sRNA detection sensitivity, prokaryotic rRNAs are often removed before preparing RNA-seq libraries (32,91,97,110,111). Successful reduction of the rRNA fraction can result in a ten-fold enrichment of the mRNA and sRNA output, and that ensures sufficient transcript coverage. Although mRNA enrichment is not absolutely necessary, it can

substantially increase transcript coverage and therefore increase the sensitivity of the experiment. The downside of enrichment might be unanticipated biases in the sequenced transcriptome. Other possibility would be to compensate the lower coverage in non-enriched samples by sequencing more cDNA, but this will significantly increase the cost of the experiment. Different methods have been used to eliminate prokaryotic rRNA, including subtractive hybridization with rRNA-specific probes, digestion with exonuclease that acts on processed transcripts, and gel electrophoresis size separation and extraction of non-rRNA fractions, which has the drawback of excluding also transcripts similar in size to rRNAs. A study compared the effectiveness and fidelity of the two most commonly used approaches, subtractive hybridization and exonuclease digestion (Figure 9), on two different synthetic metatranscriptomes and found that only subtractive hybridization adequately preserved relative transcript abundance for quantitative analyses, whereas the exonuclease treatment greatly compromised mRNA abundance fidelity (112). This is not surprising since many non-ribosomal transcripts (including sRNAs) are processed, and treatment with exonuclease results in the loss of these. Article 2 presented in this thesis describes how to use subtractive hybridization to efficiently deplete 23S, 16S and 5S rRNAs from total RNA samples.

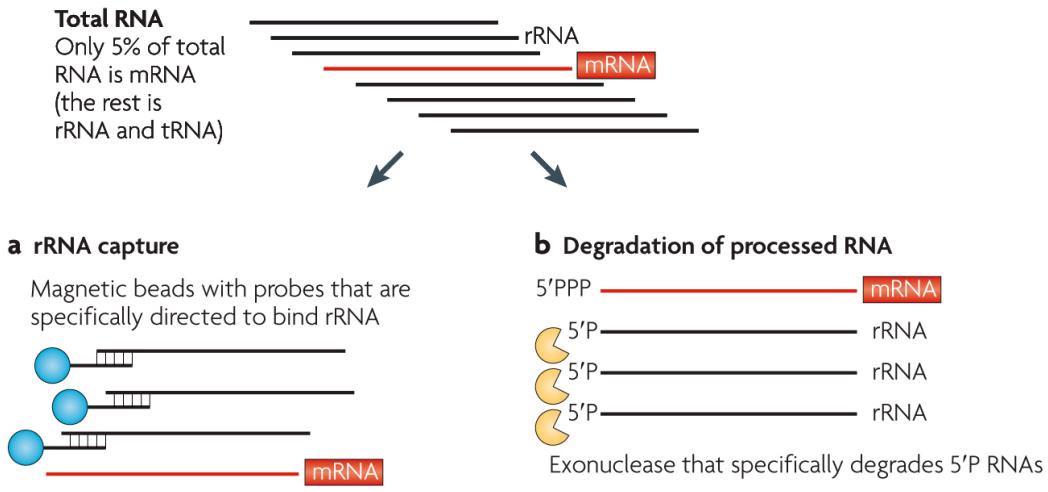


Figure 9. Methods for non-rRNA enrichment in prokaryotes (113).

Size is a key factor when detecting novel sRNAs using RNA-seq. RNA size fractionation prior to library preparation allows a greater coverage of small size RNAs and of transcripts expressed at low levels, increasing the number of novel RNAs

detected (29,32,91). As mentioned further, our study of the transcriptome of *P. aeruginosa* involved the use of three different types of sequencing libraries, with different degrees of RNA size selection (32). The libraries that contained the fraction of shorter RNAs were clearly superior in detecting sRNAs compared to a library that contained the full transcriptome, which highlights the importance of adequate library preparation to detect sRNAs (see Article 1 presented in this thesis for details). Thus, RNA-seq libraries that contain all the transcriptome will detect only a fraction of highly expressed sRNAs, unless using an enormous sequencing depth, which normally results in increased economic costs. RNA-size fractionation should be performed prior to rRNA-depletion, as the larger rRNA molecules will be excluded during size selection. The smaller rRNAs, like 5S rRNA, can be excluded after the size selection (see Article 2 presented in this thesis for details).

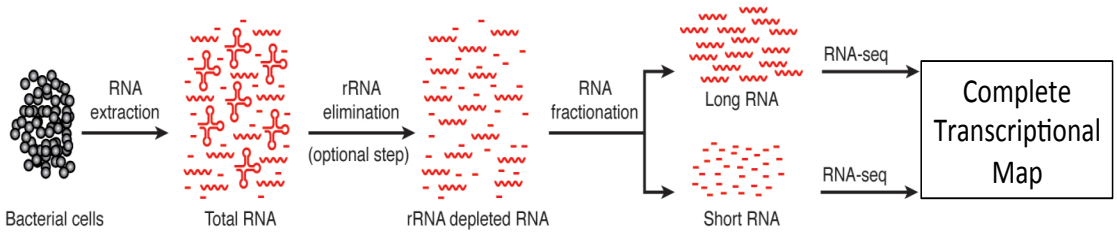


Figure 10. Summary flowchart suggesting an experimental design to define a complete transcriptional map. Figure adapted from (22).

To summarize, the sensitivity of RNA-seq for detecting sRNAs greatly depends on the design of the experiment. The most important factors when designing such experiments are:

(i) The choice of a method for RNA extraction that does retain all sRNA species. Methods as the hot-phenol protocol or the guanidinium thiocyanate-phenol-chloroform extraction allow retention of all RNA species.

(ii) Removal of abundant rRNAs increases the relative abundance of mRNA and sRNA species and ensures sufficient transcript coverage. Subtractive hybridization depletion methods are preferred over the use of exonuclease, as exonuclease digestion results in the loss of many non-ribosomal transcripts (including sRNAs) that are processed.

(iii) RNA size fractionation prior to library preparation also allows a greater coverage of small size RNAs and of transcripts expressed at low levels, and thus the sensitivity of the experiment increases.

(iv) The use of a strand-specific method for library preparation retains the directionality of RNA molecules, making possible to detect antisense transcription. Direct labeling of the RNA molecules with different adapters at the 5' and 3' ends is preferred over other methods as oligonucleotides complementary to the 5’ and 3’ adaptors are used for the synthesis of cDNA, considerably reducing the bias introduced by use of random primers to synthesize the cDNA.

4. *PSEUDOMONAS AERUGINOSA*

This Gram-negative bacillus can be found in a wide variety of environmental sources, from water to soil. It grows on a wide range of media over a wide range of temperature. It is an important opportunistic pathogen in humans, causing urinary tract infections, respiratory system infections, systemic infections in burn patients as well as being the major cause of morbidity and mortality in cystic fibrosis patients (114).

The increasing clinical relevance of *P. aeruginosa* as an opportunistic human pathogen is based in several reasons: its adaptability, its innate resistance to many antibiotics and disinfectants, its armoury of virulence factors, an increasing number of patients compromised by age, disease or immunosuppressive therapy, and its ability to persist and multiply in a great variety of moist environments and equipment in hospital wards, bathrooms and kitchens, which is of particular importance in cross-infection control. The genome of *P. aeruginosa* is one of the largest among bacteria: it contains approximately 6.2 million base pairs and 5,570 genes, including a high proportion of regulatory genes which most likely reflect its adaptability (115). The choice of *P. aeruginosa* in this study was mainly motivated by the fact that it possesses vast and complex regulatory networks, whose mechanisms are not entirely understood, and by the low number of regulatory sRNAs identified in this bacterium prior to our study (reviewed in (116)). We hypothesized that *P. aeruginosa* must have many unknown sRNAs encoded in its large genome, and that has been shown to be true.

4.1 Antibiotic resistance

P. aeruginosa is intrinsically resistant to many antimicrobial agents and acquired resistance is frequently developed during antibiotic therapy (117). The mechanisms of resistance consist of chromosomally encoded β-lactamases, multidrug efflux systems, low outer membrane permeability and lipopolysaccharide (LPS) modification. Antibiotics such as piperacillin, aztreonam, ceftazidime, meropenem,

ciprofloxacin, colistin, azithromycin or tobramycin are usually active against *P. aeruginosa*, but resistance often develops in the case of chronic infected CF patients (118-122). *P. aeruginosa* has more than 10 resistance-nodulation-division (RND) family transporters, of which 9 transport multiple antibiotics. These pumps export antibiotics but also biocides, dyes, detergents, metabolic inhibitors, organic solvents and molecules involved in bacterial cell-to-cell communication. Despite drug efflux not being the natural function of RND pumps, these pumps enable many pathogens to resist antimicrobial drugs and thus compromise treatment of infectious diseases (123). *P. aeruginosa* also produces a chromosomally encoded cephalosporinase designated AmpC. The expression of AmpC is highly inducible by certain β-lactams, such as cefoxitin, imipenem or ampicillin. The antipseudomonal penicillins (such as piperacillin) and cephalosporins (such as ceftazidime) are very weak AmpC inducers, despite the fact that this enzyme hydrolyzes them. However, during treatment with these weak inducers, mutants showing high expression of AmpC (AmpC-derepressed mutants) are frequently selected, leading to the failure of antimicrobial therapy (124). In addition *P. aeruginosa* is able to modify its outer-membrane lipopolysaccharides (LPS) leading to increased tolerance towards cationic antimicrobial peptides, like colistin, and increased inflammatory responses (125).

4.2 Quorum sensing

Quorum sensing (QS) has been defined as the chemical-mediated cell-to-cell communication systems to coordinate gene expression and group activities within bacterial communities (126). *P. aeruginosa* produces at least three small molecules that function as cell-to-cell communication signals. The acyl homoserine lactone signals 3-oxo-C_{12}-HSL and C_4-HSL have been well studied and function in combination with the LasR and RhlR proteins, respectively (127-131). Together, these two HSLs control up to 11% of the *P. aeruginosa* genes (132-134). The third *P. aeruginosa* cell-to-cell signal, named the Pseudomonas quinolone signal (PQS), is a quinolone compound that was identified as 2-heptyl-3-hydroxy-4-quinolone (135). PQS is synthesized from anthranilate by the products of the *pqs* operon *pqsABCDEH*. The precursor of the PQS molecule is 4-hydroxy-2-heptyl-quinoline (HHQ), which is released from and taken up by bacterial cells (136). The expression of *pqsABCDE* is regulated by the transcriptional regulator PqsR (also designated as MvfR) (137). By contrast, *pqsH*, which is necessary for the conversion of HHQ to

PQS, is positively regulated by LasR (138). PQS controls numerous virulence factors and it serves as a regulatory link between the *las* and *rhl* QS systems (139,140). PQS is produced in the lungs of CF patients infected with *P. aeruginosa* (141), it is required for virulence in nematodes, plants, and mice (137,138,142-144), and it also induces apoptosis and decreases viability in eukaryotic cells (145).

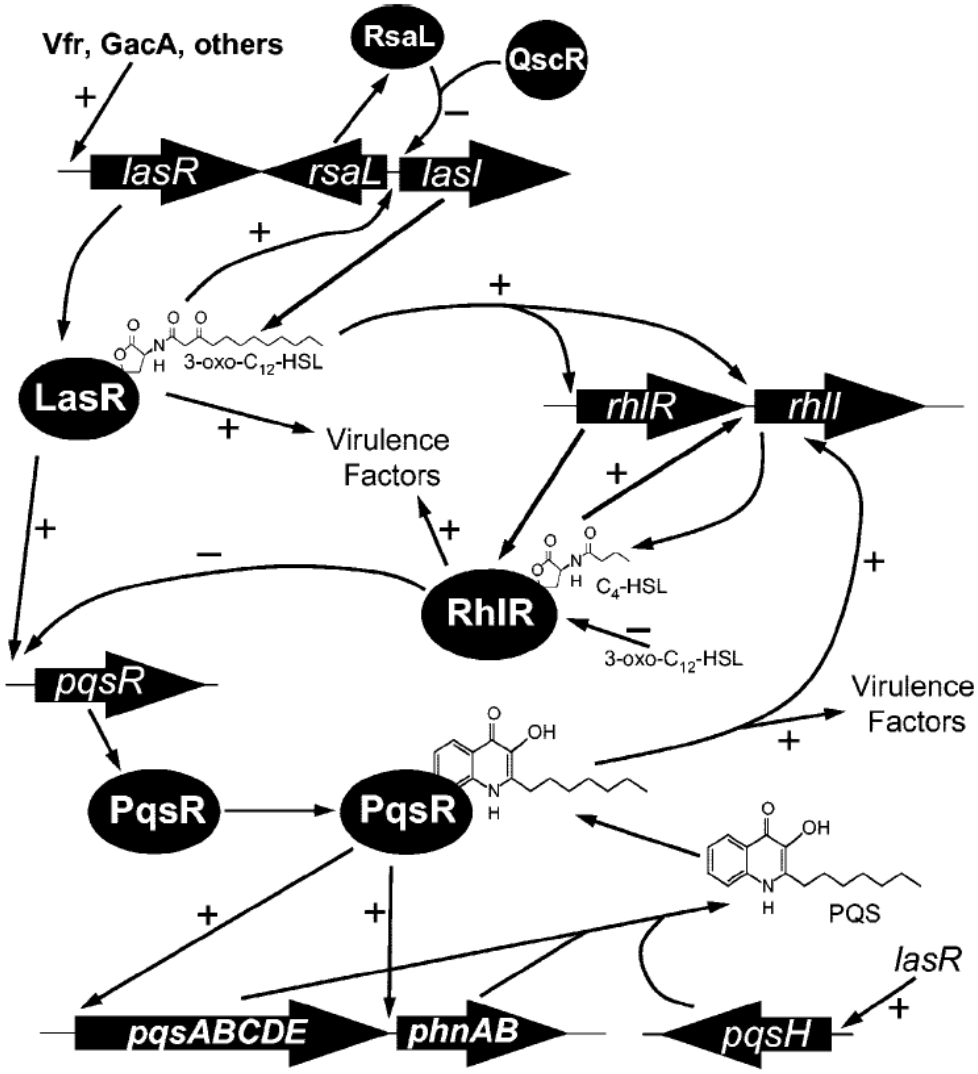


Figure 11. Model of the *P. aeruginosa* QS signaling hierarchy. Plus and minus symbols indicate positive and negative effects, respectively. The QS system consists of two HSL regulatory circuits (las and rhl) linked to a 2-alkyl-4-quinolone (PQS) system. In complex interactions with additional regulators, including Vfr, GacA, RsaL, and RpoS, the QS systems regulate the expression of several of virulence factors. Figure adapted from (146).

4.3 Small RNAs in *P. aeruginosa*

Only 44 sRNAs were identified in *P. aeruginosa* up to 2012. A variety of approaches were used to identify these 44 sRNAs in *P. aeruginosa* (reviewed in (116)) including those based on direct detection due to high abundance (147), similarity with a known sRNA of *E. coli* (148), computational predictions (149), and co-purification with a protein (60). Three systematic searches for sRNAs were undertaken in *P.aeruginosa,* where all focused on intergenic regions. In two studies sRNAs were predicted on the basis of sequence conservation, as well as the presence of rho-independent terminators, putative promoters and conserved secondary structure (85,86). A third study used a combined bioinformatics and RNomics approach, where total RNA was size-fractionated, co-immunoprecipitated with Hfq, cloned and sequenced (60).

Of the 44 sRNAs known in *P. aeruginosa*, only the role of 12 sRNAs is characterized, where 4 are classical sRNAs (6S, 4.5S, RNase P and tmRNA) (reviewed in (116)). In *P. aeruginosa*, sRNAs are known to be involved in the regulation of virulence genes (RsmY/Z), carbon catabolite repression (CrcZ), iron metabolism (PrrF1, PrrF2, PrrH), nitrogen assimilation (NrsZ), and QS regulation (PhrS) (47,78,149-152). PhrS is a 213-nt long sRNA, predicted by biocomputation (85,86) and isolated by an RNomics approach (60). PhrS is highly transcribed under low oxygen conditions, and its expression requires the oxygen-responsive regulator ANR (47) (Figure 12). PhrS activates the translation of the *pqsR* gene by base-pairing with a short untranslated open reading frame (uof) to which the *pqsR* gene is translationally coupled (47). PqsR is one of the key QS regulators in *P. aeruginosa* and its synthesis induces the PQS synthetic operon *pqsABCDE* (138,153). Beside its function in inter-bacterial signaling, PQS is required for full synthesis of the *P. aeruginosa* virulence factor pyocianin, which can react with molecular oxygen to generate oxidative stress (154). In addition to its function as a regulatory RNA, PhrS encodes a highly conserved 37 amino-acid peptide (60). The peptide contains a predicted transmembrane segment and sub-cellular fractionation revealed that the peptide is indeed located in the cytoplasmic membrane (CM). In Article 4 presented in this thesis we investigated the role of sRNA OsiS, which was first identified in the genome-wide search presented in Article 1. OsiS is highly transcribed during oxidative stress conditions, including treatment with sublethal concentration of hydrogen peroxide and lethal concentrations of ciprofloxacin. We show that by inducing the expression of OsiS the levels of the sRNA PhrS are greatly reduced. We predict that the interaction is by

direct base-pairing between the two sRNAs, with a recognition site of OsiS at the highly conserved-region of PhrS (155). However, more experiments are required to know the exact mechanism of interaction between these two sRNAs. OsiS is, to the best of our knowledge, the first sRNA whose main function seems to be regulating the cellular levels of another sRNA.

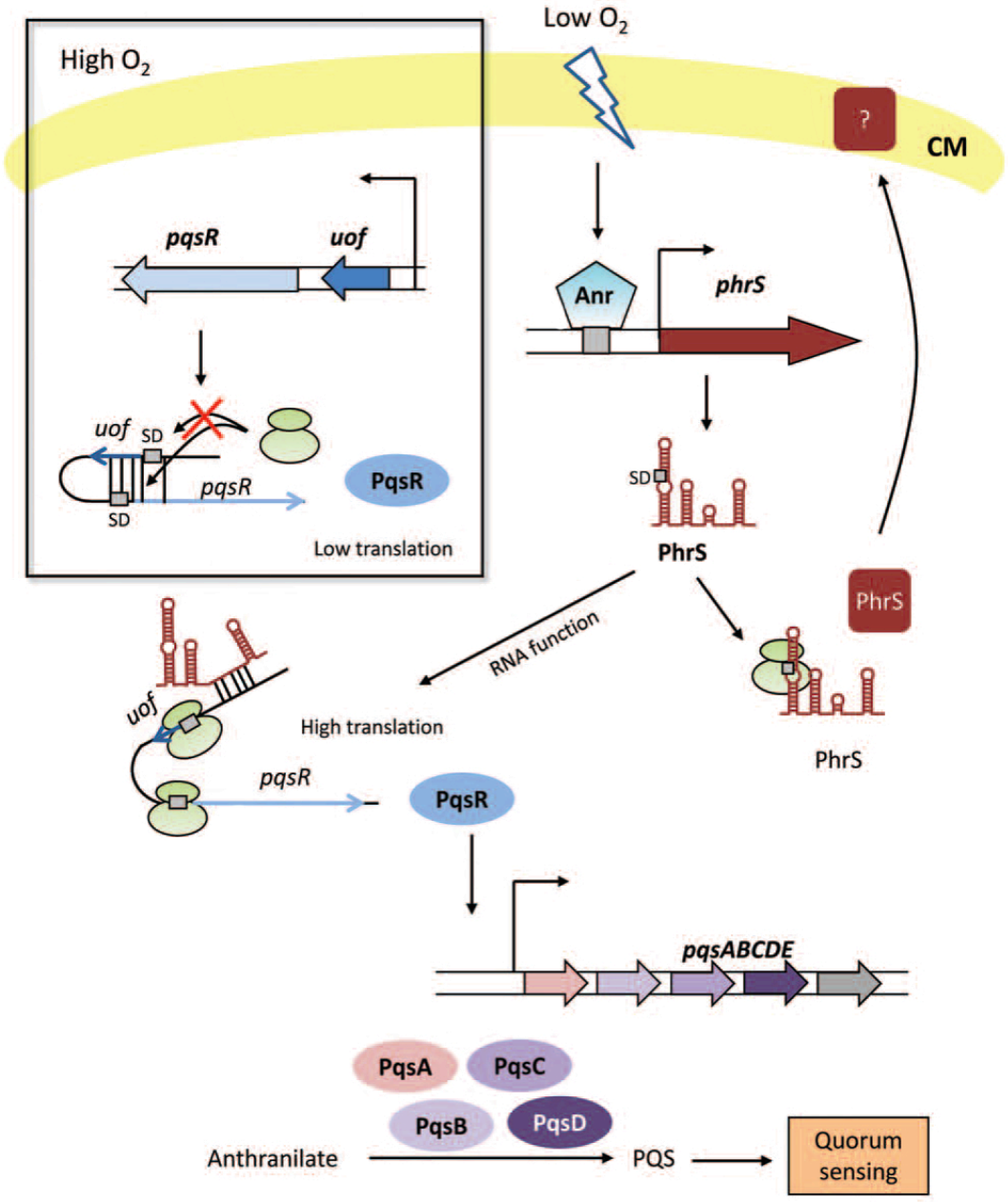


Figure 12. Quorum sensing regulation by PhrS. Figure adapted from (156).

The size and complexity of the *P. aeruginosa* genome suggested that there were many undetected sRNAs in this bacterium. In 2012, we published the first genome-wide search of sRNAs using RNA-seq in *P. aeruginosa* (32). RNA isolated from exponential and early-stationary phase cells was used to prepare three different types of libraries for sequencing. The sequencing libraries were named LIB>100, LIB<500 and LIB<200 to denote the transcript sizes in nucleotides that are included in each preparation. Library LIB>100 contained all mRNAs transcribed. Libraries LIB<500 and LIB<200 were each prepared with size selection steps and contain transcripts shorter than 500 and 200 nt respectively. We found over 500 novel intergenic sRNAs in the three libraries, where a different number of transcripts were detected in each library. The two libraries that do not include larger transcripts (LIB<500 and LIB<200) were clearly superior in detecting sRNAs. Our data suggests that the set of identified RNAs obtained depends strongly on the specific library preparation strategy used, and that RNA fractionation increases the sequencing depth favoring identification of trasncripts expressed at low copies (see Article 1 in Chapter 2 for details). In addition we identified 249 antisense sRNAs (asRNAs) in *P. aeruginosa* PAO1 grown in 13 different conditions, including osmotic and oxidative stress, early and stationary phases, and a variety of antibiotics clinically used (see Article 3 in Chapter 2 for details). In this study only one library type was used, which contained transcripts shorter than 500 nt (library LIB<500).

Later in 2012, two independent studies that used RNA-seq in *P. aeruginosa* were published. One of the studies investigated strains PAO1 and PA14 at early stationary phase using a similar preparation than that of library LIB<500. The authors found 163 novel sRNAs, of which 60 were antisense sRNAs (asRNAs) (157). Interestingly, the authors found that a number of these sRNAs were strain-specific or showed strain-specific expression (157). The other study sequenced *P. aeruginosa* PA14 grown at 28ºC and at 37ºC using a strategy similar to library LIB>100. This study identified 165 intergenic sRNAs and 384 antisense transcriptional start sites (TSS) (158). The strategy used to prepare the libraries only detected the strand specificity of TSS, thus the authors could only identify antisense TSS, but not the full length of asRNAs. The strategy used by Wurtzel *et al* (2012) enabled the identification of TSSs thanks to the differential treatment of RNA with tobacco acid pyrophosphatase (TAP). TSSs can be detected by selecting for primary RNA transcripts that have a 5′-triphosphate (processed RNAs have a 5′-monophosphate) (24). In conclusion, our sequencing approach was the most sensitive, identifying novel sRNAs, both intergenic sRNAs and full-length asRNAs, compared to the other two genome-wide studies in *P.*

aeruginosa. However our library preparations could be further improved by sequencing differential cDNA library pairs (one fraction treated with TAP or with terminator exonuclease, and the other fraction left untreated) to discriminate primary from processed 5' ends. It would be interesting to make a comparison between the three studies to assess the level of overlap between them, as well as for creating a single database with all the sRNAs detected so far in *P. aeruginosa* strains and their patterns of expression in the conditions studied.

5. CONCLUSION AND PERSPECTIVES

Identifying the full set of transcripts, including large and small RNAs, is crucial for a comprehensive study of the transcriptome. Until recently, our knowledge of the transcriptome was largely derived from gene predictions based on sequence conservation and limited evidence, and has thus been partial. During the last decade, RNA-seq has revealed the complex landscape and dynamics of the transcriptome of all organisms studied at an unprecedented level of sensitivity and accuracy. Compared with traditional low-throughput Sanger sequencing technology, which only detects the more abundant transcripts, the enormous sequencing depth of a typical RNA-seq experiment can offer an almost complete snapshot of a transcriptome, if using appropriate library preparations. As mentioned earlier, the main focus of this thesis has been to investigate the number and characteristics of sRNAs in the highly adaptable pathogen *P. aeruginosa*. RNA-seq protocols have been optimized in order to increase their sensitivity and detect as many sRNAs as possible. The research articles written as part of this work are enclosed in the next section. **Article 1** describes how a combination of three different RNA-seq libraries is used to identify novel transcripts encoded in intergenic regions. Almost all known sRNAs and over 500 novel intergenic sRNAs were identified with this approach. The use of three libraries in combination increased the number of novel transcripts identified. However there were significant differences in the subset of transcripts detected in each library, underscoring the importance of library preparation strategy and relative sRNA abundance for successful sRNA detection. **Article 2** is a book chapter that provides detailed experimental procedures for preparing the three different sequencing libraries used in Article 1, as well as a computational pipeline to analyze the data generated and find both intergenic and asRNAs. One of the libraries (LIB>100) is prepared using a standard RNA-seq protocol for full transcriptomes and contains information on all mRNAs transcribed by the bacteria. The other two libraries (LIB<500 and LIB<200) are prepared from size-selected RNA containing transcripts shorter than 500 and 200 nt, respectively. The different libraries can either be used alone or in combination to increase the number of identified sRNAs. In **Article 3** 249 asRNAs are identified using the library type LIB<500. *P. aeruginosa* PAO1 was grown in 13 different conditions to ensure a comprehensive characterization of the *P. aeruginosa* asRNA-ome. The conditions studied include exponential and early-stationary phase, osmotic, oxidative and antibiotic stress. A considerable number of

asRNAs were transcribed opposite to genes encoding membrane proteins and genes involved in the transport of small molecules, cell wall, lipopolysaccharide (LPS), and capsule biosynthesis. A substantial number of asRNAs significantly changed their expression under osmotic, oxidative and antibiotic stress, suggesting that asRNAs play a regulatory role during these conditions. Finally, **Article 4** investigates the role of sRNA OsiS that was first identified in the genome-wide search presented in Article 1. OsiS is highly transcribed during oxidative stress conditions. We show that by inducing the expression of OsiS the levels of the sRNA PhrS are greatly reduced. PhrS activates the translation of PqsR under low oxygen concentrations, which in turn activates the synthesis of PQS. Thus, OsiS links the oxygen levels to the production of QS molecules. OsiS is another example of a sRNA that responds to oxygen levels in bacteria, such as PhrS in *P. aeruginosa* (47), RgsA in *P. fluorescens* (86), OxyS, FnrS and MicF in *E. coli* (159-162), and RliB in *L. monocytogenes* (15). OsiS is, to the best of our knowledge, the first sRNA whose main function seems to be regulating the cellular levels of another sRNA.

The study presented in Article 1 is the first to use RNA-seq to discover novel sRNAs in *P. aeruginosa* and utilized three types of sequencing libraries to maximize the number of detected sRNAs. Over 500 novel sRNA transcripts were identified, increasing the number of known sRNAs in *P. aeruginosa* by a factor of 12 and revealing a large set of potential regulatory molecules in an already highly regulated species. Two libraries that do not include larger transcripts (LIB<500 and LIB<200) were clearly superior in detecting sRNAs. Although the use of three libraries increased the number of novel transcripts identified, there were significant differences in the subset of transcripts detected in each library, underscoring the importance of library preparation strategy and relative sRNA abundance for successful sRNA detection. The limited extent of sRNA sequence conservation observed among pseudomonads highlights the challenges of sRNA bioinformatic prediction and emphasizes the importance of sRNAs at the species level in monitoring changes in environmental conditions. The approach described can be applied to identify sRNAs in any bacterium under different growth and stress conditions.

The multitude of sRNAs found in Articles 1 and 3 suggests the involvement of sRNAs in probably all the regulatory networks in *P. aeruginosa*, and characterizing their role will be highly important for our understanding of an organism with such a versatile life style that can successfully move from the environment into the human body as a

pathogen. There are many questions that remain to be answered regarding the physiological roles of bacterial sRNAs. Even with the recent advances in transcriptomics, the hundreds of sRNAs now being reported in all bacteria studied need to be further validated and functionally characterized before we can really grasp the importance/role of bacterial sRNAs.

Interestingly, few of the known bacterial sRNAs genes have been identified by mutational screens. This might be due to the fact that most of the sRNA genes do not code for proteins and are thus immune to frameshift or nonsense mutations. In addition, sRNAs are often very small (few hundred nucleotides) and sometimes multicopy, which makes them difficult (even impossible) targets for mutational screens. This could also be a consequence of the regulatory rather than essential nature of many sRNAs. Marvig *et al* (2013) studied the transmissible *P. aeruginosa* DK2 clone isolated from Danish CF patients, and analyzed the genomes of 55 bacterial isolates collected from 21 infected individuals over 38 years (163). Subsequent to transmission, sub-lineages of DK2 evolved independently for years in separate patients, enabling the study of parallel evolution and identification of genes targeted by mutations to increase pathogen fitness in the CF airway. The authors found that the genes targeted by mutations were related to antibiotic resistance, cell envelope, or regulatory functions, and that the prevalence of adaptive mutations correlated with the evolutionary success of the sub-lineages. We looked for conserved adaptive mutations in all the intergenic sRNAs in *P. aeruginosa*, the known sRNAs prior to our genome-wide study plus the 513 novel sRNAs detected by it. We could not find any conserved adaptive mutation in any of the sRNA genes, suggesting that though sRNAs play important roles in regulatory networks, they are not the primary targets of bacterial evolution. In agreement with this, a long-term evolution experiment with *E. coli*, still ongoing, has not yet revealed any relevant changes in sRNA genes (164). However, a short-term evolution experiment reported a spontaneous mutation in *Myxococcus xanthus* that eliminated the regulatory function of the Pxr RNA, affecting multicellular fruiting body development (165). It will be interesting to perform more evolution experiments in other species to assess the role of sRNA regulation in the long-term adaptation of bacteria.

From our work with OsiS, we infer that *P. aeruginosa* most likely needs a tight and fine-tuned regulation of the mechanisms that allow adapting to shifts in oxygen tension, and that the sRNA OsiS is involved in this kind of regulation by responding to high oxygen tension and regulating the levels of sRNA PhrS. PhrS is assumed to counteract the low levels of PQS under hypoxic conditions, when PQS production is limited (47). The sRNA OsiS role seems to "counteract the counteracting effect" of sRNA PhrS at high oxygen tension. So both sRNAs OsiS and PhrS respond to oxygen levels and allow the fine-tuning of PQS synthesis. In laboratory conditions, the growth rate of mutants deficient in PhrS production is the same as the wild-type strain, though the deletion mutant produced less pyocianin. We could not produce a strain lacking the OsiS gene, however over-expression of OsiS does not affect any apparent phenotype. Does that mean that sRNAs PhrS and OsiS have a minor regulatory role? And what is the biological implication of oxygen-dependent fine-tuning of PQS? During infection, *P. aeruginosa* is subject to constant shifts in oxygen tension due to the human airway structure, biofilm formation, antibiotic treatment and the host immune response. Thus, fine-tuning regulators as PhrS and OsiS make no extraordinary difference in controlled and steady laboratory conditions, however we hypothesized that they might be pivotal for a fast and fine-tuned regulation of the mechanisms necessary to survive during adaptation to an environment with such fluctuating levels of oxygen tension as the one inside the host.

As the number of sRNAs detected increases day by day, it would be interesting to understand why RNA-based regulation is sometimes preferred over protein-based regulation. It has been suggested previously that sRNA regulation has benefits over using protein regulators. These benefits might include reduced metabolic cost, additional levels of regulation, faster regulation and unique regulatory properties (reviewed in (166)). The coupled degradation of both the regulatory sRNA and the target mRNAs allows to achieve responses that are different than those obtained with transcriptional factors (167). However, it seems that in many cases a combination of both types regulation seems to be the way in which bacteria regulate the cellular levels of certain proteins and RNAs. For example, PhrS RNA regulates the expression of the *pqsR* gene (47), which is also regulated by other transcriptional factors, such as LasR and RhlR (146). In this case the regulation by transcriptional factors and PhrS is counteractive, as PhrS promotes the expression of *pqsR* in conditions when the transcriptional regulators limit its expression. Similarly, PhrS is regulated by both the oxygen-responsive regulator ANR and by OsiS RNA (Article 4). In this case the regulation of both protein and sRNA act in a complementary way,

ANR activating the expression of PhrS under low oxygen conditions, and OsiS reducing the levels of PhrS under high oxygen conditions. Thus, regulating a target at two levels, with both a transcription regulator and base-pairing sRNA, allows modifying the dynamics of target gene expression (as the regulation of *pqsR* by transcriptional factors and PhrS RNA) as well as reducing leakiness (as the regulation of PhrS by ANR and OsiS RNA).

As we start to solve some of the traditional challenges for studying sRNAs, such as the lack of technical skills to detect them, new challenges arise. One of them is the handling of RNA-seq data. Comparing these extremely large data sets between different studies is difficult, as library preparation protocols, sequencing platforms, and thresholds for detecting transcripts differ from study to study. It is necessary to address considerations like what is an appropriate threshold of reads for transcript detection, or how reproducibility should be reported. Ironically, we now have such an enormous volume of data that it is hard to use them. Universally available platforms are already and will be needed for sharing, analyzing, and storing transcriptomic data. An even greater challenge than sharing and comparing RNA-seq data, will be the validation and characterization of the hundreds of novel sRNAs reported by every new study. While Northern blot analysis has been the golden standard for validating RNAs, its sensitivity is significantly lower than that achieved by RNA-seq. Techniques such as 5' and 3' RACE experiments might be a good alternative for validating sRNAs. Just as Northern blots, RACE experiments provide information about the size and processing of a transcript. However, contrary to Northern analysis, RACE experiments provide single-base resolution of the coordinates of the transcripts. Without higher throughput techniques than the now available, it will take years, probably decades, to characterize the hundreds of sRNAs now being reported. The studies that characterize the function of sRNAs are mostly based in deletion strains and sRNA overexpression (168-170). Unfortunately, the phenotypes associated with decreased or increased expression of sRNAs are often subtle and might only be detected under specific conditions, as is the case of strains overexpressing PhrS and OsiS RNAs. The understanding of the molecular mechanisms of action of individual novel RNAs will most likely be one of the biggest challenges of current RNA research.

REFERENCES

1. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009 Feb 20;136(4):615–28.

2. Crick F. Central dogma of molecular biology. Nature. 1970 Aug 8;227(5258):561–3.

3. Gesteland RF, Cech TR, Atkins JF. The RNA World. CSHL Press. 2006.

4. Brachet J, Chantrenne H. The function of the nucleus in the synthesis of cytoplasmic proteins. Cold Spring Harbor symposia on quantitative biology. 1956;21:329–37.

5. Crick FH. On protein synthesis. Symposia of the Society for Experimental Biology. 1958.

6. Hoagland MB, Stephenson ML, Scott JF, et al. A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem. 1958 Mar;231(1):241–57.

7. Brenner S, Jacob F, Meselson M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature. 1961 May 13;190:576–81.

8. Faust PLP, Kornfeld SS, Chirgwin JMJ. Cloning and sequence analysis of cDNA for human cathepsin D. Proceedings of the National Academy of Sciences. 1985 Jul 31;82(15):4910–4.

9. Schena MM, Shalon DD, Davis RWR, Brown POP. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995 Oct 19;270(5235):467–70.

10. Selinger DWD, Cheung KJK, Mei RR, Johansson EME, Richmond CSC, Blattner FRF, et al. RNA expression analysis using a 30 base pair resolution Escherichia coli genome array. Nat Biotechnol. 2000 Nov 30;18(12):1262–8.

11. Merrell DS, Butler SM, Qadri F, Dolganov NA, Alam A, Cohen MB, et al. Host-induced epidemic spread of the cholera bacterium. Nature. 2002 Jun 5;417(6889):642–5.

12. Revel ATA, Talaat AMA, Norgard MVM. DNA microarray analysis of differential gene expression in Borrelia burgdorferi, the Lyme disease spirochete. Proceedings of the National Academy of Sciences. 2002 Feb 4;99(3):1562–7.

13. Belland RJR, Zhong GG, Crane DDD, Hogan DD, Sturdevant DD, Sharma JJ, et al. Genomic transcriptional profiling of the developmental cycle of Chlamydia trachomatis. Proceedings of the National Academy of Sciences. 2003 Jul 7;100(14):8478–83.

14. Mäurer APA, Mehlitz AA, Mollenkopf HJH, Meyer TFT. Gene expression profiles of Chlamydophila pneumoniae during the developmental cycle and iron depletion-mediated persistence. PLoS Pathog. 2007 May 31;3(6):e83–3.

15. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The Listeria transcriptional landscape from saprophytism to virulence. Nature. 2009;459(7249):950–6.

16. Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JCD. Unravelling the biology

of macrophage infection by gene expression profiling of intracellular Salmonella enterica. Molecular Microbiology. 2002 Dec 31;47(1):103–18.

17. Perocchi F, Xu Z, Clauder-Münster S, Steinmetz LM. Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. Nucleic Acids Research. 2006 Dec 31;35(19):e128–8.

18. Canales RDR, Luo YY, Willey JCJ, Austermiller BB, Barbacioru CCC, Boysen CC, et al. Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotechnol. 2006 Aug 31;24(9):1115–22.

19. Westermann AJ, Gorski SA, Vogel J. Dual RNA-seq of pathogen and host. Nat Rev Micro. 2012 Sep;10(9):618–30.

20. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57–63.

21. Martin JAJ, Wang ZZ. Next-generation transcriptome assembly. Nat Rev Genet. 2011 Sep 30;12(10):671–82.

22. Lasa I, Toledo-Arana A, Gingeras TR. An effort to make sense of antisense transcription in bacteria. RNA Biol. 2012 Aug;9(8):1039–44.

23. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Current Opinion in Microbiology. 2007 Apr;10(2):102–9.

24. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature. 2010 Mar 11;464(7286):250–5.

25. Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, Grills G, et al. Transcriptome Analysis of Pseudomonas syringae Identifies New Genes, Noncoding RNAs, and Antisense Activity. Journal of Bacteriology. 2010 May 1;192(9):2359–72.

26. Guell M, Van Noort V, Yus E, Chen W-H, Leigh-Bell J, Michalodimitrakis K, et al. Transcriptome Complexity in a Genome-Reduced Bacterium. Science. 2009 Nov 27;326(5957):1268–71.

27. Dornenburg JE, DeVita AM, Palumbo MJ, Wade JT. Widespread antisense transcription in Escherichia coli. mBio. 2010 Apr;1(1).

28. Albrecht M, Sharma CM, Reinhardt R, Vogel J, Rudel T. Deep sequencing-based discovery of the Chlamydia trachomatis transcriptome. Nucleic Acids Research. 2010 Jan;38(3):868–77.

29. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, et al. Genome-wide antisense transcription drives mRNA processing in bacteria. Proc Natl Acad Sci USA. 2011 Dec 13;108(50):20172–7.

30. Rasmussen S, Nielsen HB, Jarmer H. The transcriptionally active regions in the genome of Bacillus subtilis. Molecular Microbiology. Wiley Online Library; 2009;73(6):1043–57.

31. Parkhomchuk DD, Borodina TT, Amstislavskiy VV, Banaru MM, Hallen LL, Krobitsch SS, et al. Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Research. 2009 Sep 30;37(18):e123–3.

32. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Genome-wide identification of novel small RNAs in Pseudomonas aeruginosa. Environmental Microbiology. 2012 Aug;14(8):2006–16.

33. Stougaard P, Molin S, Nordström K. RNAs involved in copy-number control and incompatibility of plasmid R1. Proceedings of the National Academy of Sciences. 1981 Oct;78(10):6008–12.

34. Tomizawa J, Itoh T, Selzer G, Som T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proceedings of the National Academy of Sciences. 1981 Mar;78(3):1421–5.

35. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993 Dec 3;75(5):843–54.

36. Ambros V. MicroRNAs and developmental timing. Current opinion in genetics & development. 2011 Aug;21(4):511–7.

37. Ambros V. The functions of animal microRNAs. Nature. Nature Publishing Group; 2004 Sep 16;431(7006):350–5.

38. Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2011 Dec;3(12):a00379.

39. Repoila F, Darfeuille F. Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. Biology of the Cell. 2009;101:117–31.

40. Breaker RR. Prospects for riboswitch discovery and analysis. Molecular Cell. 2011 Sep 16;43(6):867–79.

41. Karginov FV, Hannon GJ. The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea. Molecular Cell. Elsevier Ltd; 2010 Jan 15;37(1):7–19.

42. Vanderpool CK, Balasubramanian D, Lloyd CR. Dual-function RNA regulators in bacteria. Biochimie. 2011 Nov;93(11):1943–9.

43. Holmqvist E. Macromolecular Matchmaking: Mechanisms and Biology of Bacterial Small RNAs. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and technology. 2012 Apr 3;:916.

44. Mizuno T, Chou MY, Inouye M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proceedings of the National Academy of Sciences. 1984 Apr;81(7):1966–70.

45. Sharma CM, Darfeuille F, Plantinga TH, Vogel J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes & development. 2007 Nov 1;21(21):2804–17.

46. Večerek B, Moll I, Bläsi U. Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. EMBO J. 2007 Feb 1;26(4):965–75.

47. Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, et al. The small RNA PhrS stimulates synthesis of the Pseudomonas aeruginosa quinolone signal. Molecular Microbiology. 2011 May;80(4):868–85.

48. Urban JH, Vogel J. Two seemingly homologous noncoding RNAs act hierarchically to activate glmS mRNA translation. PLoS Biol. 2008 Mar 18;6(3):e64.

49. Hammer BK, Bassler BL. Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic Vibrio cholerae. Proceedings of the National Academy of Sciences. 2007 Jul 3;104(27):11145–9.

50. Carpousis AJ, Luisi BF, McDowall KJ. Endonucleolytic Initiation of mRNA Decay in Escherichia coli. Progress in Molecular Biology and Translational Science. Progress in Molecular Biology and Translational Science; 2009. pp. 91–135.

51. Celesnik H, Deana A, Belasco JG. Initiation of RNA Decay in Escherichia coli by 5′ Pyrophosphate Removal. Molecular Cell. Elsevier; 2007;27(1):79–90.

52. Obana N, Shirahama Y, Abe K, Nakamura K. Stabilization of Clostridium perfringens collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence. Molecular Microbiology. 2010 Sep;77(6):1416–28.

53. Ramirez-Peña E, Treviño J, Liu Z, Perez N, Sumby P. The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. Molecular Microbiology. 2010 Dec;78(6):1332–47.

54. Kawamoto H, Koide Y, Morita T, Aiba H. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. Molecular Microbiology. 2006 Aug;61(4):1013–22.

55. Vogel J, Ben F Luisi. Hfq and its constellation of RNA. Nat Rev Micro. Nature Publishing Group; 2011 Aug 1;9(8):578–89.

56. Hopkins JF, Panja S, Woodson SA. Rapid binding and release of Hfq from ternary complexes during RNA annealing. Nucleic Acids Research. 2011 Jul;39(12):5193–202.

57. Hwang W, Arluison V, Hohng S. Dynamic competition of DsrA and rpoS fragments for the proximal binding site of Hfq as a means for efficient annealing. Nucleic Acids Research. 2011 Jul;39(12):5131–9.

58. Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA. Positive regulation by small RNAs and the role of Hfq. Proc Natl Acad Sci USA. 2010 May 25;107(21):9602–7.

59. Maki K, Morita T, Otaka H, Aiba H. A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. Molecular Microbiology. 2010 May;76(3):782–92.

60. Sonnleitner E, Sorger-Domenigg T, Madej MJ, Findeiss S, Hackermüller J, Hüttenhofer A, et al. Detection of small RNAs in Pseudomonas aeruginosa by RNomics and structure-based bioinformatic tools. Microbiology (Reading, Engl). 2008 Oct;154(Pt 10):3175–87.

61. Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Sogaard-Andersen L, Kallipolitis BH. Identification of small Hfq-binding RNAs in Listeria monocytogenes. RNA. 2006 Jul;12(7):1383–96.

62. Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes & development. 2001 Jul 1;15(13):1637–51.

63. Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S. Global analysis of small RNA and mRNA targets of Hfq. Molecular Microbiology. 2003 Nov;50(4):1111–24.

64. Chaulk SG, Smith Frieday MN, Arthur DC, Culham DE, Edwards RA, Soo P, et al. ProQ is an RNA chaperone that controls ProP levels in Escherichia coli. Biochemistry. 2011 Apr 19;50(15):3095–106.

65. Pandey SP, Minesinger BK, Kumar J, Walker GC. A highly conserved protein of unknown function in Sinorhizobium meliloti affects sRNA regulation similar to Hfq. Nucleic Acids Research. 2011 Jun;39(11):4691–708.

66. Thomason MK, Storz G. Bacterial Antisense RNAs: How Many Are There, and What Are They Doing? *. Annu Rev Genet. 2010 Dec;44(1):167–88.

67. Callen BP, Shearwin KE, Egan JB. Transcriptional Interference between Convergent Promoters Caused by Elongation over the Promoter. Molecular Cell. 2004 Jun 3;14(5):10–0.

68. Blomberg PP, Wagner EGE, Nordström KK. Control of replication of plasmid R1: the duplex between the antisense RNA, CopA, and its target, CopT, is processed specifically in vivo and in vitro by RNase III. EMBO J. 1990 Jun 30;9(7):2331–40.

69. Gerdes KK, Nielsen AA, Thorsted PP, Wagner EGE. Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable hok, srnB and pndA effector messenger RNAs. Journal of Molecular Biology. 1992 Aug 4;226(3):637–49.

70. Lee E-J, Groisman EA. An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. Molecular Microbiology. 2010 May;76(4):1020–33.

71. Silby MW, Levy SB. Overlapping protein-encoding genes in Pseudomonas fluorescens Pf0-1. PLoS Genet. 2008 Jun;4(6):e1000094.

72. Sayed N, Jousselin A, Felden B. A cis-antisense RNA acts in trans in Staphylococcus aureus to control translation of a human cytolytic peptide. Nat Struct Mol Biol. 2012 Jan;19(1):105–12.

73. Arnvig KB, Young DB. Identification of small RNAs in Mycobacterium tuberculosis. Molecular Microbiology. 2009 Aug;73(3):397–408.

74. Mandin P, Gottesman S. Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. EMBO J. 2010 Sep 15;29(18):3094–107.

75. Papenfort K, Said N, Welsink T, Lucchini S, Hinton JCD, Vogel J. Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. Molecular Microbiology. 2009 Oct;74(1):139–58.

76. Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P. Identification of new noncoding RNAs in Listeria monocytogenes and prediction of mRNA targets. Nucleic Acids Research. 2007;35(3):962–74.

77. Babitzke P, Romeo T. CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Current Opinion in Microbiology. 2007 Apr;10(2):156–63.

78. Sonnleitner E, Abdou L, Haas D. Small RNA as global regulator of carbon catabolite repression in Pseudomonas aeruginosa. Proc Natl Acad Sci USA. 2009 Dec 22;106(51):21866–71.

79. Wassarman KM. 6S RNA: a regulator of transcription. Molecular Microbiology. 2007 Sep;65(6):1425–31.

80. Neusser T, Polen T, Geissen R, Wagner R. Depletion of the non-coding regulatory 6S RNA in E. coli causes a surprising reduction in the expression of the translation machinery. BMC Genomics. 2010;11:165.

81. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding

frontiers. Molecular Cell. 2011 Sep 16;43(6):880–91.

82. Vogel J, Sharma CM. How to find small non-coding RNAs in bacteria. Biol Chem. 2005 Dec;386(12):1219–38.

83. Pánek J, Bobek J, Mikulík K, Basler M, Vohradský J. Biocomputational prediction of small non-coding RNAs in Streptomyces. BMC Genomics. 2008;9:217.

84. Xiao B, Li W, Guo G, Li B, Liu Z, Jia K, et al. Identification of small noncoding RNAs in Helicobacter pylori by a bioinformatics-based approach. Current microbiology. 2009 Mar;58(3):258–63.

85. Livny J, Brencic A, Lory S, Waldor M. Identification of 17 Pseudomonas aeruginosa sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic Acids Research. 2006;34(12):3484.

86. Gonzalez N, Heeb S, Valverde C, Kay E, Reimmann C, Junier T, et al. Genome-wide search reveals a novel GacA-regulated small RNA in Pseudomonas species. BMC Genomics. 2008;9:167.

87. Wassarman KM, Zhang A, Storz G. Small RNAs in Escherichia coli. Trends in Microbiology. 1999 Jan;7(1):37–45.

88. Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner E, Margalit H, et al. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Current Biology. 2001;11(12):941–50.

89. Rivas E, Klein RJ, Jones TA, Eddy SR. Computational identification of noncoding RNAs in E. coli by comparative genomics. Curr Biol. 2001 Sep 4;11(17):1369–73.

90. Altuvia S. Identification of bacterial small non-coding RNAs: experimental approaches. Current Opinion in Microbiology. 2007;10(3):257–61.

91. Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, Camilli A. Experimental discovery of sRNAs in Vibrio cholerae by direct cloning, 5S/tRNA depletion and parallel sequencing. Nucleic Acids Research. 2009 Apr 1;37(6):e46.

92. Irnov I, Sharma CM, Vogel J, Winkler WC. Identification of regulatory RNAs in Bacillus subtilis. Nucleic Acids Research. 2010 Oct;38(19):6637–51.

93. Weissenmayer BA, Prendergast JGD, Lohan AJ, Loftus BJ. Sequencing Illustrates the Transcriptional Response of Legionella pneumophila during Infection and Identifies Seventy Novel Small Non-Coding RNAs. PLoS ONE. 2011 Mar 3;6(3):e17570.

94. Oliver HF, Orsi RH, Ponnala L, Keich U, Wang W, Sun Q, et al. Deep RNA sequencing of L. monocytogenes reveals overlapping and extensive stationary phase and sigma B-dependent transcriptomes, including multiple highly transcribed noncoding RNAs. BMC Genomics. 2009;10:641.

95. Mraheil MA, Billion A, Mohamed W, Mukherjee K, Kuenne C, Pischimarov J, et al. The intracellular sRNA transcriptome of Listeria monocytogenes during growth in macrophages. Nucleic Acids Research. 2011 May;39(10):4235–48.

96. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, et al. Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq. PLoS Genet. 2008 Aug 22;4(8):e1000163.

97. Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, et al. A Strand-Specific RNA–Seq Analysis of the Transcriptome of the Typhoid Bacillus Salmonella

Typhi. PLoS Genet. 2009 Jul 17;5(7):e1000569.

98. Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, et al. The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. Proc Natl Acad Sci USA. 2012 May 15;109(20):E1277–86.

99. Vockenhuber M-P, Sharma CM, Statt MG, Schmidt D, Xu Z, Dietrich S, et al. Deep sequencing-based identification of small non-coding RNAs in Streptomyces coelicolor. RNA Biol. 2011 May 1;8(3).

100. Mitschke J, Georg J, Scholz I, Sharma CM, Dienst D, Bantscheff J, et al. An experimentally anchored map of transcriptional start sites in the model cyanobacterium Synechocystis sp. PCC6803. Proc Natl Acad Sci USA. 2011 Feb 1;108(5):2124–9.

101. Bradley ES, Bodi K, Ismail AM, Camilli A. A Genome-Wide Approach to Discovery of Small RNAs Involved in Regulation of Virulence in Vibrio cholerae. PLoS Pathog. 2011 Jul 14;7(7):e1002126.

102. Koo JT, Alleyne TM, Schiano CA, Jafari N, Lathem WW. Global discovery of small RNAs in Yersinia pseudotuberculosis identifies Yersinia-specific small, noncoding RNAs required for virulence. Proc Natl Acad Sci USA. 2011 Aug 29.

103. Raghavan R, Groisman EA, Ochman H. Genome-wide detection of novel regulatory RNAs in E. coli. Genome Research. 2011 Sep;21(9):1487–97.

104. Shinhara A, Matsui M, Hiraoka K, Nomura W, Hirano R, Nakahigashi K, et al. Deep sequencing reveals as-yet-undiscovered small RNAs in Escherichia coli. BMC Genomics. 2011 Aug 24;12(1):428.

105. Arnvig KB, Comas I, Thomson NR, Houghton J, Boshoff HI, Croucher NJ, et al. Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. PLoS Pathog. 2011 Nov;7(11):e1002342.

106. Wilms I, Overlöper A, Nowrousian M, Sharma CM, Narberhaus F. Deep sequencing uncovers numerous small RNAs on all four replicons of the plant pathogen Agrobacterium tumefaciens. RNA Biol. 2012 Apr;9(4):446–57.

107. Karpinets TV, Greenwood DJ, Sams CE, Ammons JT. RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. BMC Biol. 2006;4:30.

108. Zhao J, Hyman L, Moore C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiology and Molecular Biology Reviews. 1999 Jun;63(2):405–45.

109. Urich T, Lanzén A, Qi J, Huson DH, Schleper C, Schuster SC. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. PLoS ONE. 2008;3(6):e2527.

110. Passalacqua KD, Varadarajan A, Ondov BD, Okou DT, Zwick ME, Bergman NH. Structure and complexity of a bacterial transcriptome. Journal of Bacteriology. 2009 May;191(10):3203–11.

111. Yoder-Himes DR, Chain PSG, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, et al. Mapping the Burkholderia cenocepacia niche response via high-throughput sequencing. Proc Natl Acad Sci USA. 2009 Mar 10;106(10):3976–81.

112. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG, et al. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. Nat Meth. 2010 Oct;7(10):807–12.

113. Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nat Rev Genet. 2010;11(1):9–16.

114. Brennan AL, Geddes DM. Cystic fibrosis. Curr Opin Infect Dis. 2002 Apr;15(2):175–82.

115. Stover CKC, Pham XQX, Erwin ALA, Mizoguchi SDS, Warrener PP, Hickey MJM, et al. Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature. 2000 Aug 30;406(6799):959–64.

116. Sonnleitner E, González N, Haas D. Small RNAs of Pseudomonas spp. Pseudomonas. 2010;:3–28.

117. Hancock REW, Speert DP. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and impact on treatment. Drug Resist Updat. 2000 Aug;3(4):247–55.

118. Saiman L, Marshall BC, Mayer-Hamblett N, Burns JL, Quittner AL, Cibene DA, et al. Azithromycin in patients with cystic fibrosis chronically infected with Pseudomonas aeruginosa: a randomized controlled trial. JAMA. 2003 Oct 1;290(13):1749–56.

119. Doring G, Conway SP, Heijerman HG, Hodson ME, Hoiby N, Smyth A, et al. Antibiotic therapy against Pseudomonas aeruginosa in cystic fibrosis: a European consensus. 2000. pp. 749–67.

120. Aronoff SC, Klinger JD. In vitro activities of aztreonam, piperacillin, and ticarcillin combined with amikacin against amikacin-resistant Pseudomonas aeruginosa and P. cepacia isolates from children with cystic fibrosis. Antimicrobial agents and chemotherapy. Am Soc Microbiol; 1984;25(2):279–80.

121. Szaff M, Hoiby N, Flensborg EW. Frequent antibiotic therapy improves survival of cystic fibrosis patients with chronic Pseudomonas aeruginosa infection. Acta Paediatr Scand. 1983 Sep;72(5):651–7.

122. Høiby N. Recent advances in the treatment of Pseudomonas aeruginosa infections in cystic fibrosis. BMC Medicine. BioMed Central Ltd; 2011 Apr 4;9(1):32.

123. Poole K. Bacterial multidrug efflux pumps serve other functions. Microbe-American Society For Microbiology. ASM PRESS; 2008;3(4):179.

124. Livermore D. Beta-lactamases in laboratory and clinical resistance. Clinical Microbiology Reviews. 1995;8(4):557.

125. Ernst R, Yi E, Guo L, Lim K, Burns J, Hackett M, et al. Specific lipopolysaccharide found in cystic fibrosis airway Pseudomonas aeruginosa. Science. 1999;286(5444):1561.

126. Fuqua WC, Winans SC, Greenberg EP. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. Journal of Bacteriology. 1994 Jan;176(2):269–75.

127. Gambello MJ, Iglewski BH. Cloning and characterization of the Pseudomonas aeruginosa lasR gene, a transcriptional activator of elastase expression. Journal of Bacteriology. 1991 May;173(9):3000–9.

128. Ochsner UA, Koch AK, Fiechter A, Reiser J. Isolation and characterization of a regulatory gene affecting rhamnolipid biosurfactant synthesis in Pseudomonas

aeruginosa. Journal of Bacteriology. 1994 Apr;176(7):2044–54.

129. Ochsner UA, Reiser J. Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences. 1995 Jul 3;92(14):6424–8.

130. Pearson JP, Gray KM, Passador L, Tucker KD, Eberhard A, Iglewski BH, et al. Structure of the autoinducer required for expression of Pseudomonas aeruginosa virulence genes. Proceedings of the National Academy of Sciences. 1994 Jan 4;91(1):197–201.

131. Pearson JP, Passador L, Iglewski BH, Greenberg EP. A second N-acylhomoserine lactone signal produced by Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences. 1995 Feb 28;92(5):1490–4.

132. Schuster M, Lostroh CP, Ogi T, Greenberg EP. Identification, timing, and signal specificity of Pseudomonas aeruginosa quorum-controlled genes: a transcriptome analysis. Journal of Bacteriology. 2003 Apr;185(7):2066–79.

133. Wagner VE, Bushnell D, Passador L, Brooks AI, Iglewski BH. Microarray analysis of Pseudomonas aeruginosa quorum-sensing regulons: effects of growth phase and environment. Journal of Bacteriology. 2003 Apr;185(7):2080–95.

134. Whiteley M, Lee KM, Greenberg EP. Identification of genes controlled by quorum sensing in Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences. 1999 Nov 23;96(24):13904–9.

135. Pesci EC, Milbank JB, Pearson JP, McKnight S, Kende AS, Greenberg EP, et al. Quinolone signaling in the cell-to-cell communication system of Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences. 1999 Sep 28;96(20):11229–34.

136. Déziel E, Lepine F, Milot S, He J, Mindrinos MN, Tompkins RG, et al. Analysis of Pseudomonas aeruginosa 4-hydroxy-2-alkylquinolines (HAQs) reveals a role for 4-hydroxy-2-heptylquinoline in cell-to-cell communication. Proceedings of the National Academy of Sciences. 2004 Feb 3;101(5):1339–44.

137. Cao HH, Krishnan GG, Goumnerov BB, Tsongalis JJ, Tompkins RR, Rahme LGL. A quorum sensing-associated virulence gene of Pseudomonas aeruginosa encodes a LysR-like transcription regulator with a unique self-regulatory mechanism. Proceedings of the National Academy of Sciences. 2001 Dec 3;98(25):14613–8.

138. Gallagher LA, McKnight SL, Kuznetsova MS, Pesci EC, Manoil C. Functions required for extracellular quinolone signaling by Pseudomonas aeruginosa. Journal of Bacteriology. 2002 Dec;184(23):6472–80.

139. Diggle SP, Winzer K, Chhabra SR, Worrall KE, Cámara M, Williams P. The Pseudomonas aeruginosa quinolone signal molecule overcomes the cell density-dependency of the quorum sensing hierarchy, regulates rhl-dependent genes at the onset of stationary phase and can be produced in the absence of LasR. Molecular Microbiology. 2003 Aug 11;50(1):29–43.

140. McKnight SL, Iglewski BH, Pesci EC. The Pseudomonas quinolone signal regulates rhl quorum sensing in Pseudomonas aeruginosa. Journal of Bacteriology. 2000 May;182(10):2702–8.

141. Collier DN, Anderson L, McKnight SL, Noah TL, Knowles M, Boucher R, et al. A bacterial cell to cell signal in the lungs of cystic fibrosis patients. FEMS Microbiology Letters. 2002 Sep 24;215(1):41–6.

142. Lau GW, Ran H, Kong F, Hassett DJ, Mavrodi D. Pseudomonas aeruginosa pyocyanin is critical for lung infection in mice. Infection and immunity. 2004 Jul;72(7):4275–8.

143. Mahajan-Miklos S, Tan MW, Rahme LG, Ausubel FM. Molecular mechanisms of bacterial virulence elucidated using a Pseudomonas aeruginosa-Caenorhabditis elegans pathogenesis model. Cell. 1999 Jan 8;96(1):47–56.

144. Rahme LG, Tan MW, Le L, Wong SM, Tompkins RG, Calderwood SB, et al. Use of model plant hosts to identify Pseudomonas aeruginosa virulence factors. Proceedings of the National Academy of Sciences. 1997 Nov 25;94(24):13245–50.

145. Calfee MW, Shelton JG, McCubrey JA, Pesci EC. Solubility and bioactivity of the Pseudomonas quinolone signal are increased by a Pseudomonas aeruginosa-produced surfactant. Infection and immunity. 2005 Feb;73(2):878–82.

146. Wade DS, Calfee MW, Rocha ER, Ling EA, Engstrom E, Coleman JP, et al. Regulation of Pseudomonas quinolone signal synthesis in Pseudomonas aeruginosa. Journal of Bacteriology. 2005 Jul;187(13):4372–80.

147. Vogel DW, Hartmann RK, Struck JCR, Ulbrich N, Erdmann VA. The sequence of the 6S RNA gene of Pseudomonas aeruginosa. Nucleic Acids Research. 1987;15(11):4583–91.

148. Toschka HY, Struck JC, Erdmann VA. The 4.5S RNA gene from Pseudomonas aeruginosa. Nucleic Acids Research. 1989 Jan 11;17(1):31–6.

149. Wilderman PJ, Sowa NA, FitzGerald DJ, FitzGerald PC, Gottesman S, Ochsner UA, et al. Identification of tandem duplicate regulatory small RNAs in Pseudomonas aeruginosa involved in iron homeostasis. Proc Natl Acad Sci USA. 2004 Jun 29;101(26):9792–7.

150. Brencic A, McFarland KA, McManus HR. The GacS/GacA signal transduction system of Pseudomonas aeruginosaacts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. Molecular Microbiology. 2009 Aug;73(3):434–45.

151. Oglesby-Sherrouse A, Vasil M. Characterization of a heme-regulated non-coding RNA encoded by the prrF locus of Pseudomonas aeruginosa. PLoS ONE. 2010;5(4):e9930.

152. Wenner N, Maes A, Cotado-Sampayo M, Lapouge K. NrsZ: a novel, processed, nitrogen-dependent small non-coding RNA that2 regulates Pseudomonas aeruginosa PAO1 virulence. Environmental Microbiology. 2013;:Availableonlineaheadofprint.

153. DArgenio DA, Calfee MW, Rainey PB, Pesci EC. Autolysis and autoaggregation in Pseudomonas aeruginosa colony morphology mutants. Journal of Bacteriology. 2002 Dec;184(23):6481–9.

154. Price-Whelan A, Dietrich LEP, Newman DK. Rethinking “secondary” metabolism: physiological roles for phenazine antibiotics. Nat Chem Biol. 2006 Feb;2(2):71–8.

155. Gómez-Lozano M, Molin S. Small RNA OsiS links oxidative stress to quorum sensing control in. Manuscript in preparation. 2013 Sep 6;:1–17.

156. Sonnleitner E, Romeo A, Bläsi U. Small regulatory RNAs in Pseudomonas aeruginosa. RNA Biol. 2012 Apr;9(4):364–71.

157. Ferrara S, Brugnoli M, De Bonis A, Righetti F, Delvillani F, Dehò G, et al.

Comparative Profiling of Pseudomonas aeruginosa Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs. PLoS ONE. 2012;7(5):e36553.

158. Wurtzel O, Yoder-Himes DR, Han K, Dandekar AA, Edelheit S, Greenberg EP, et al. The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. PLoS Pathog. 2012 Sep;8(9):e1002945.

159. Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. Cell. 1997 Jul 11;90(1):43–53.

160. Boysen A, Møller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. J Biol Chem. 2010 Apr 2;285(14):10690–702.

161. Durand S, Storz G. Reprogramming of anaerobic metabolism by the FnrS small RNA. Molecular Microbiology. Wiley Online Library; 2010;75(5):1215–31.

162. Chou JH, Greenberg JT, Demple B. Posttranscriptional repression of Escherichia coli OmpF protein in response to redox stress: positive control of the micF antisense RNA by the soxRS locus. Journal of Bacteriology. 1993 Feb;175(4):1026–31.

163. Marvig RL, Johansen HK, Molin S, Jelsbak L. Genome Analysis of a Transmissible Lineage of Pseudomonas aeruginosa Reveals Pathoadaptive Mutations and Distinct Evolutionary Paths of Hypermutators. PLoS Genet. 2013 Sep;9(9):e1003741.

164. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009 Oct 28;461(7268):1243–7.

165. Yu Y-TN, Yuan X, Velicer GJ. Adaptive evolution of an sRNA that controls Myxococcus development. Science. 2010 May 21;328(5981):993.

166. Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiology Reviews. Wiley Online Library; 2010;34(5):866–82.

167. Levine EE, Zhang ZZ, Kuhlman TT, Hwa TT. Quantitative characteristics of gene regulation by small RNA. Audio, Transactions of the IRE Professional Group on. 2007 Aug 31;5(9):e229–9.

168. Hobbs ECE, Astarita JLJ, Storz GG. Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. Audio, Transactions of the IRE Professional Group on. 2009 Dec 31;192(1):59–67.

169. Massé E, Vanderpool CK, Gottesman S. Effect of RyhB small RNA on global iron use in Escherichia coli. Journal of Bacteriology. Am Soc Microbiol; 2005;187(20):6962–71.

170. Papenfort K, Pfeiffer V, Lucchini S, Sonawane A, Hinton JCD, Vogel J. Systematic deletion of Salmonella small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. Molecular Microbiology. 2008 Apr 30;68(4):890–906.

6. RESEARCH ARTICLES

This section consists of the full-length research articles that were prepared as part of this PhD project. The articles are enclosed in the following order:

Article 1

Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*.

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S. (2012). *Environmental Microbiology*, 14(8), 2006–2016. doi:10.1111/j.1462-2920.2012.02759.x

Article 2

Identification of bacterial small RNAs by RNA sequencing.

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S. (2013). *Methods in Pseudomonas aeruginosa. Accepted for publication.*

Article 3

Antisense small RNAs respond to osmotic, oxidative and antibiotic stress in *Pseudomonas aeruginosa*.

Gómez-Lozano, M., Marvig, R. L., Tulstrup, M. V. L., Tribelli, P., & Molin, S. (2013). *Manuscript in preparation.*

Article 4

Small RNA OsiS links oxidative stress to quorum sensing control in *Pseudomonas aeruginosa*.

Gómez-Lozano, M. & Molin, S. (2013). *Manuscript in preparation.*

Article 1

Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*.

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S.
Environmental Microbiology, **2012**, 14(8), 2006–2016.
doi:10.1111/j.1462-2920.2012.02759.x

environmental microbiology

Environmental Microbiology (2012) **14**(8), 2006–2016 doi:10.1111/j.1462-2920.2012.02759.x

Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*

María Gómez-Lozano,[1,2] Rasmus Lykke Marvig,[1] Søren Molin[1,2] and Katherine S. Long[1,2]*

[1]*Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.*

[2]*Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.*

Summary

Bacterial small regulatory RNAs (sRNAs) function in post-transcriptional control of gene expression and control a variety of processes including metabolic reactions, stress responses and pathogenesis in response to environmental signals. A variety of approaches have been used previously to identify 44 sRNAs in the opportunistic human pathogen *Pseudomonas aeruginosa*. In this work, RNA sequencing (RNA-seq) is used to identify novel transcripts in *P. aeruginosa* involving a combination of three different sequencing libraries. Almost all known sRNAs and over 500 novel intergenic sRNAs are identified with this approach. Although the use of three libraries increased the number of novel transcripts identified, there were significant differences in the subset of transcripts detected in each library, underscoring the importance of library preparation strategy and relative sRNA abundance for successful sRNA detection. Nearly 90% of the novel sRNAs have no orthologous bacterial sequences outside of *P. aeruginosa*, supporting a limited degree of sequence conservation and rapid evolution of sRNAs at the species level. We anticipate that the data will be useful for the study of regulatory sRNAs in bacteria and that the approach described here may be applied to identify sRNAs in any bacterium under different growth and stress conditions.

Introduction

Bacteria contain numerous small regulatory RNAs (sRNAs) ranging in size from 70 to 500 nucleotides (nt) that modulate gene expression (Waters and Storz, 2009; Gottesman and Storz, 2011). Most sRNAs function through base-pairing interactions with messenger RNA (mRNA) targets that can affect translation initiation, transcription termination or mRNA stability. Others function by binding to and altering the activity of proteins that regulate gene expression. Small RNAs are known to control diverse adaptation processes, where bacterial physiology is adjusted in response to environmental cues. There are examples of sRNAs that are involved in transcription reprogramming, carbon metabolism, iron homeostasis, cell envelope homeostasis and coordination of virulence (Repoila and Darfeuille, 2009). Although most current information on sRNA function is derived from *Escherichia coli*, progress is being made in the functional characterization of sRNAs in other bacteria (Papenfort and Vogel, 2010). In the important human pathogen *Pseudomonas aeruginosa*, there is knowledge on the roles of 12 out of the 44 sRNAs identified thus far, where four are classical sRNAs (6S, 4.5S, RNase P and tmRNAs) and the others function as regulators of primary and secondary metabolism (Sonnleitner and Haas, 2011).

A variety of approaches have been used to identify individual sRNAs in *P. aeruginosa* (reviewed in Sonnleitner *et al*., 2010) including those based on direct detection due to high abundance (Vogel *et al*., 1987), similarity with a known sRNA of *E. coli* (Toschka *et al*., 1989), computational predictions (Wilderman *et al*., 2004) and copurification with a protein (Sonnleitner *et al*., 2008). Three systematic searches for sRNAs have been undertaken in *P. aeruginosa*, where all have focused on intergenic regions. In two studies sRNAs were predicted on the basis of sequence conservation, as well as the presence of rho-independent terminators, putative promoters and conserved secondary structure. Although both studies used similar computational approaches followed by experimental verification by Northern blotting, two largely distinct sets of novel sRNAs were identified (Livny *et al*., 2006; González *et al*., 2008). A third study used a combined bioinformatics and RNomics approach, where total RNA was size-fractionated, co-immunoprecipitated with Hfq, cloned and sequenced (Sonnleitner *et al*., 2008).

Recent genome-wide studies to identify sRNAs have been largely based on tiling arrays and next-generation sequencing technologies. The latter approach, in particular, has revolutionized sRNA discovery by enabling interrogation of the transcriptome at unprecedented depths. Deep sequencing has revealed hundreds of previously

Received 15 February, 2012; revised 27 March, 2012; accepted 30 March, 2012. *For correspondence. E-mail kalon@biosustain.dtu.dk; Tel. (+45) 45 25 80 24; Fax (+45) 45 25 80 01.

undetected transcripts in diverse bacteria including *Bacillus subtilis* (Irnov *et al*., 2010), *Chlamydia trachomatis* (Albrecht *et al*., 2010), *Helicobacter pylori* (Sharma *et al*., 2010), *Legionella pneumophila* (Weissenmayer *et al*., 2011), *Listeria monocytogenes* (Oliver *et al*., 2009; Mraheil *et al*., 2011), *Salmonella enterica* (Sittka *et al*., 2008; Perkins *et al*., 2009), *Streptomyces coelicolor (Vockenhuber et al., 2011), Synechocystis* (Mitschke *et al*., 2011), *Vibrio cholerae* (Liu *et al*., 2009; Bradley *et al*., 2011) and *Yersinia pseudotuberculosis* (Koo *et al*., 2011). The detected transcripts derive from reads that map to both intergenic and coding regions, and in the sense and antisense directions relative to annotated genes. The many newly reported transcripts have uncovered new concepts on bacterial transcriptomes, including the presence of pervasive antisense transcription (Croucher and Thomson, 2010), but also raise issues of how to define sRNAs and the extent to which detected transcripts are functional (Storz *et al*., 2011).

The size and complexity of the *P. aeruginosa* genome suggests that it encodes many hitherto undetected sRNAs. In this study, we have used high-throughput cDNA sequencing (RNA-seq) to identify sRNAs in *P. aeruginosa*. RNA isolated from exponential and early-stationary phase cells was used to prepare three different types of libraries for sequencing. We find a total of over 500 novel intergenic sRNAs in the three libraries, where a different number of transcripts are detected in each library. As a significant number of unique sRNAs are detected in each library, the data suggest that the set of identified RNAs obtained depends strongly on the specific library preparation strategy used.

Results and discussion

Experimental approach for global sRNA identification

In order to optimize sRNA detection, three different types of sequencing libraries were generated using RNA isolated from *P. aeruginosa* grown to exponential and early-stationary phase in broth culture at 37°C (Fig. 1, Table 1). All steps including cell growth, RNA isolation and library preparation were performed in triplicate for each growth condition and sequenced using the Illumina HiSeq2000 platform. The sequencing libraries are named LIB>100, LIB<500 and LIB<200 to denote the transcript sizes in nucleotides that are not excluded from each, but do not imply that all transcripts in these size ranges are necessarily detected in the corresponding library. The LIB>100 library contains all mRNAs transcribed by the bacteria and was prepared using a standard RNA-seq protocol for full transcriptomes. This library does not contain sRNA transcripts shorter than 100 nt because the corresponding cDNA fragments are not retained during the protocol. The LIB<500 and LIB<200 libraries were each prepared with size selection steps and contain transcripts shorter than 500 and 200 nt respectively. These two libraries were prepared by ligating specific adapters to the 5′ and 3′ ends of the transcripts before cDNA synthesis, enabling retention of fragments corresponding to transcripts shorter than 100 nt. They are therefore strand-specific, where the information about the strand from which the sRNA is transcribed is retained. In addition, the RNA used to prepare the LIB<500 and LIB<200 libraries was treated with Tobacco Acid Pyrophosphatase to ensure the inclusion of primary transcripts containing a 5′-triphosphate. Although primary and processed transcripts are not distinguishable in these libraries, it is possible to modify the approach with additional libraries to obtain this information. The library preparation protocols also differ in terms of whether they include fragmentation steps. The protocols for preparation of the LIB>100 and LIB<500 libraries include fragmentation steps using divalent cations and RNase III respectively, whereas preparation of the LIB<200 library does not involve a fragmentation step due to the short lengths of the transcripts. Thus, library LIB<200 does not have any biases associated with fragmentation.

On average, 16.7 million sequence reads were generated from each library, and of these 12.3 million were of sufficient quality to be mapped onto the *P. aeruginosa* PAO1 genome (Table S1). As ribosomal RNAs (rRNAs) account for the vast majority of total RNA in a bacterial cell (Karpinets *et al*., 2006), these RNAs were removed by subtractive hybridization to enrich samples for sRNAs. Although depletion of 16S and 23S rRNAs using this approach is common (He *et al*., 2010), the 5S rRNA was also depleted in this work using a total of five oligonucleotides complementary to the *P. aeruginosa* rRNA sequences (Fig. S1, Table S2). The average percentage of read coverage corresponding to 23S, 16S and 5S rRNAs was 7.1%, 3.8% and 0.15% (Table S1) respectively, indicating that the rRNA depletion procedure functioned satisfactorily in the case of 23S and 16S rRNAs and efficiently in the case of 5S rRNA. A previous study used an oligonucleotide-directed RNase H digestion approach to deplete 5S rRNA and transfer RNAs (tRNAs), involving 4 and 25 oligonucleotides respectively (Liu *et al*., 2009). Although no attempt was made to deplete tRNAs in this work, the average percentage of read coverage corresponding to tRNAs was 2.7% (Table S1), indicating that these RNAs represent only a minor fraction of the total reads in non-depleted samples, and that their targeted removal is not necessary.

Identification of previously annotated sRNAs

Forty-four sRNAs have been identified previously in *P. aeruginosa* (reviewed in Sonnleitner and Haas, 2011)

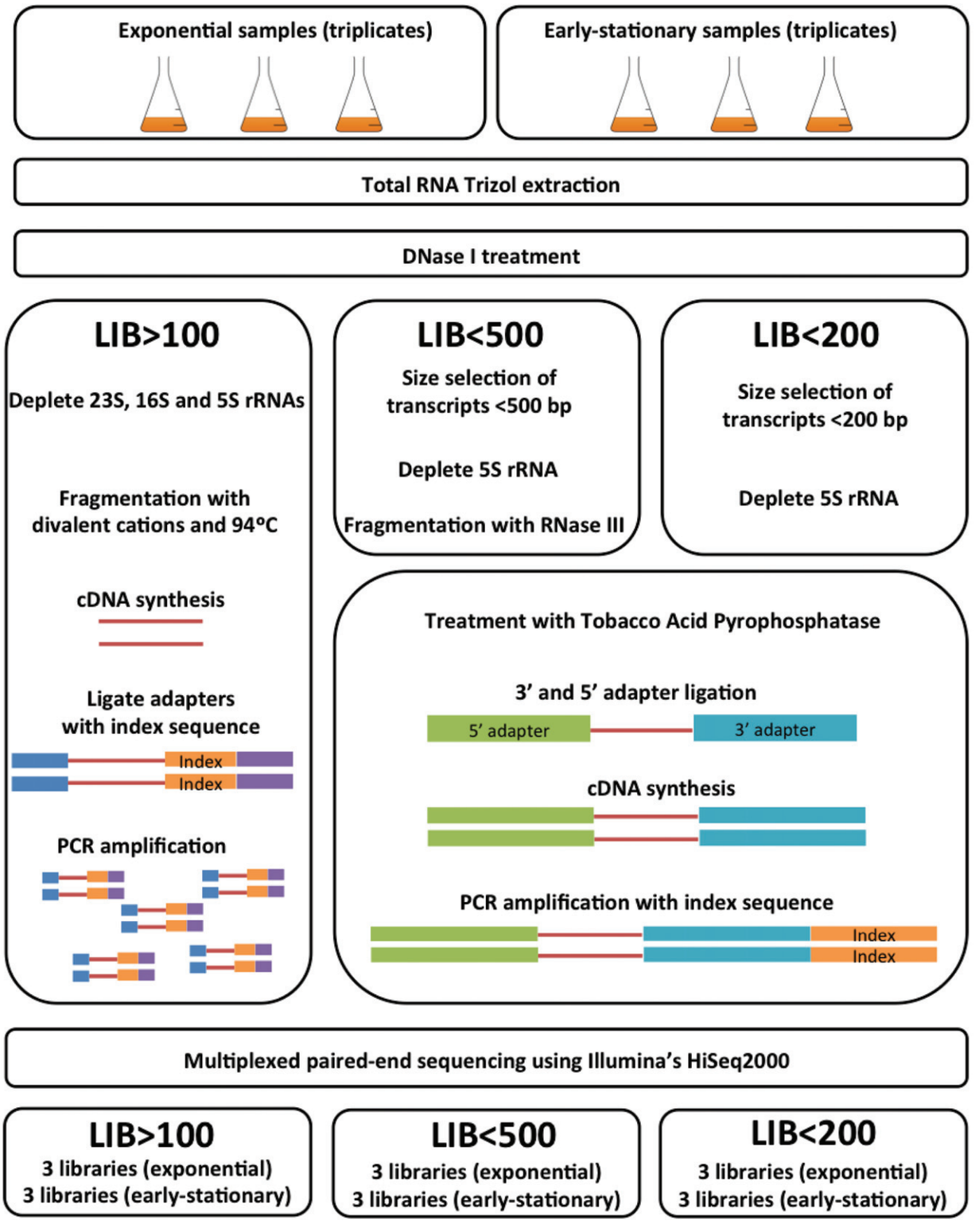


Fig. 1. Library preparation scheme. Three different sequencing libraries were generated (LIB>100, LIB<500 and LIB<200) using RNA isolated from *P. aeruginosa* cells grown to exponential and early-stationary phase in broth culture at 37°C. Each library was prepared using triplicates for each condition.

Table 1. Library characteristics.

	Library LIB>100	Library LIB<500	Library LIB<200
Length of transcripts sequenced	>100 nt	<500 nt	<200 nt
Fragmentation	Divalent cations	RNase III	No
Strand specificity retained	No	Yes	Yes

(Table S3). In order to test the accuracy and specificity of the sequencing approaches used here, the expression of the already identified sRNAs was inspected visually. All of the sRNAs except one could be visually identified in at least one of the three libraries. The sRNA that could not be identified, PA2744.1 (P13), was predicted and validated by Northern blotting in an earlier study, but the transcript length was unclear (Livny *et al*., 2006). The present data indicate that there is a 190 nt sRNA encoded in the same intergenic region on the same strand as PA2744.1. This sRNA partially overlaps with PA2744.1, but the coordinates do not correspond to those predicted previously (Table S3). This suggests that the coordinates of PA2744.1 may be 3106752–3106942.

Visual inspection of RNA-Seq data is not convenient for a genome-wide search of novel transcripts, where a systematic and reproducible method must be used. We have developed a custom-made script designed to detect intergenic sRNAs (see *Experimental procedures*). We checked how many already known sRNAs our systematic script would detect. Of the 44 already reported sRNAs, four would not be identified by the script (P26, PA4451.1, P32, PA5316.1). The reason for this is that the peaks corresponding to these sRNAs do not start and/or terminate within an intergenic region and these are not detected by the script to prevent misclassification of untranslated regions (UTRs) as ncRNAs. The sRNAs detected in each library type using the script are shown in Fig. 2. A total of 20, 34 and 36 sRNAs were detected in the LIB>100, LIB<200 and LIB<500 libraries respectively, where 18 were detected in all three libraries and 33 were detected in more than one library (Fig. 2). Thus, 39 out of 44 already annotated sRNAs could be detected in only two experimental conditions with the RNA-seq approach used here. The results, in particular those from the LIB<500 and LIB<200 libraries, are a validation of the sensitivity of RNA-seq at detecting sRNAs.

Identification of novel intergenic sRNAs

As the specific focus of this work was to identify novel intergenic sRNAs, a script was designed to analyse transcripts from all of the intergenic regions of *P. aeruginosa* PAO1 (see *Experimental procedures*). The script only detects transcripts that begin and end within an intergenic region in order to minimize misclassification of UTRs as sRNAs. A total of 513 novel sRNAs were identified by this approach and were named pant (for *P. aeruginosa* novel transcript) followed by an identification number (Table S4). The set of novel transcripts may include those corresponding to hitherto un-annotated genes and dual-function sRNAs coding for small proteins. It is also possible that a subset of the RNAs identified here derive from 5′ or 3′ UTRs. However, this does not preclude that these RNAs are functionally important or could function as sRNAs, as there is precedence for *cis*-acting RNA elements functioning as *trans*-acting RNAs. This suggests that the regulatory function of a particular RNA sequence is context-dependent and the division between *cis*- and *trans*-acting RNAs is not always sharp. For instance, the fate of riboswitch RNAs that function by transcription termination is unknown, but evidence in the literature suggests that their stability and half-lives in the cell vary (reviewed in Bastet *et al*., 2011). A specific example is that of SAM riboswitches in *L. monocytogenes* that function as *trans*-acting sRNAs to regulate the expression of the virulence regulator PrfA (Loh *et al*., 2009).

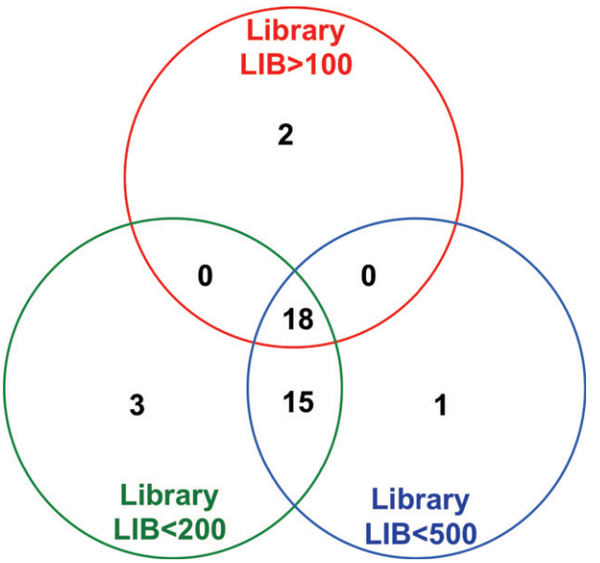


Fig. 2. Venn diagram summarizing the already annotated sRNAs detected in the three different libraries.

The plethora of novel sRNAs identified in this work raises the issue of how many of these are functional and play regulatory roles in the cell. Determination of the fraction of functional transcripts requires the characterization and detailed analysis of each sRNA. Although functional assignment may be challenging and expression of some RNAs may only be induced under specific conditions, it is likely that the majority of novel transcripts represent functional sRNAs. New and unexpected functional roles for RNA are continuously being discovered. This includes

RNAs that are relatively well understood, as recently in the case of mammalian tRNAs that, in addition to their well-known housekeeping roles in translation, can also function as regulators of gene expression (Rudinger-Thirion *et al*., 2011).

Although many novel transcripts are identified, not all transcripts are detected in all three libraries and there are significant differences in the set of transcripts detected in each library. A total of 63, 415 and 412 novel transcripts are detected in the LIB>100, LIB<500 and LIB<200 libraries respectively, indicating that significantly more transcripts are detected in the latter two libraries compared with the first (Fig. 3A). This is consistent with the fact that the LIB>100 library contains the full transcriptome, and hence a lower relative percentage of sRNAs compared with the other libraries. The lengths of the detected sRNAs range from 70 to 1117 nt with an average length of 188 nt. The length distribution of detected sRNAs in each library is shown in Fig. 3B. The lengths of the sRNAs detected in the LIB<500 and LIB<200 libraries are similar and for the most part shorter than 200 nt, whereas a significant fraction of those detected in the LIB>100 library are longer than 200 nt.

A significant overlap in sRNAs detected between the libraries is observed (Fig. 3A) with 326 novel sRNAs being identified in more than one library. A more surprising result is that 36% of the novel sRNAs were only detected in one library. Most of these novel sRNAs were detected either in the LIB<200 or LIB<500 libraries, but the length distributions of the novel RNAs detected in these libraries are nearly identical showing that these RNAs are not uniquely detected on the basis of length (data not shown). Rather, the basis of unique detection is likely due to the different library preparation strategies, in particular the inclusion of size selection and fragmentation steps. Moreover, there is an inverse relationship between the size range of RNA molecules represented in a library and its sensitivity in detecting transcripts expressed at low levels. Thus, the LIB<200 library is more sensitive than the other two libraries, and the LIB<500 library is more sensitive than the LIB>100 library. Although this work is centred on intergenic sRNAs, the LIB<500 and LIB<200 libraries can also be used to identify *cis*-encoded sRNAs because they are strand-specific and lack larger transcripts that could mask sRNAs.

Some different examples of sRNA detection in the three libraries are illustrated in Fig. 4. The already annotated sRNA PA1112.1 is only detected in the LIB>100 library, but based on its size (300 nt) it is also expected to be detected in the LIB<500 library. The probable explanation for this is the different fragmentation methods used in preparation of the LIB>100 and LIB<500 libraries, where transcripts sensitive to RNase III cleavage will not be detected in the latter. The pant154 transcript is only detected in the LIB<500 library (Fig. 4). It is not detected in the LIB<200 and LIB>100 libraries due to its size (234 nt) and relative low abundance relative to mRNA respectively. The pant55 transcript is only detected in the LIB<200 library, probably due to its small size (89 nt) and low abundance. However, it should be noted that there is also a weak signal for pant55 in the LIB<500 library that is below the cut-off value, indicating that sRNAs with a low

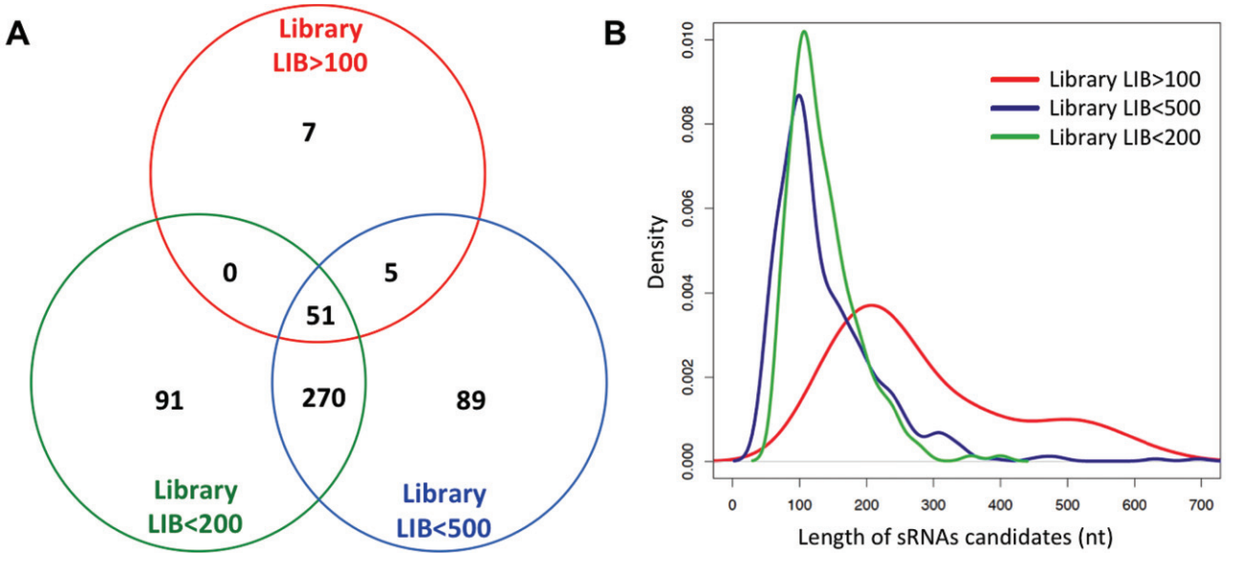


Fig. 3. A. Venn diagram illustrating the novel sRNAs detected in the three different libraries.
B. Length distributions of the sRNAs detected in each library.

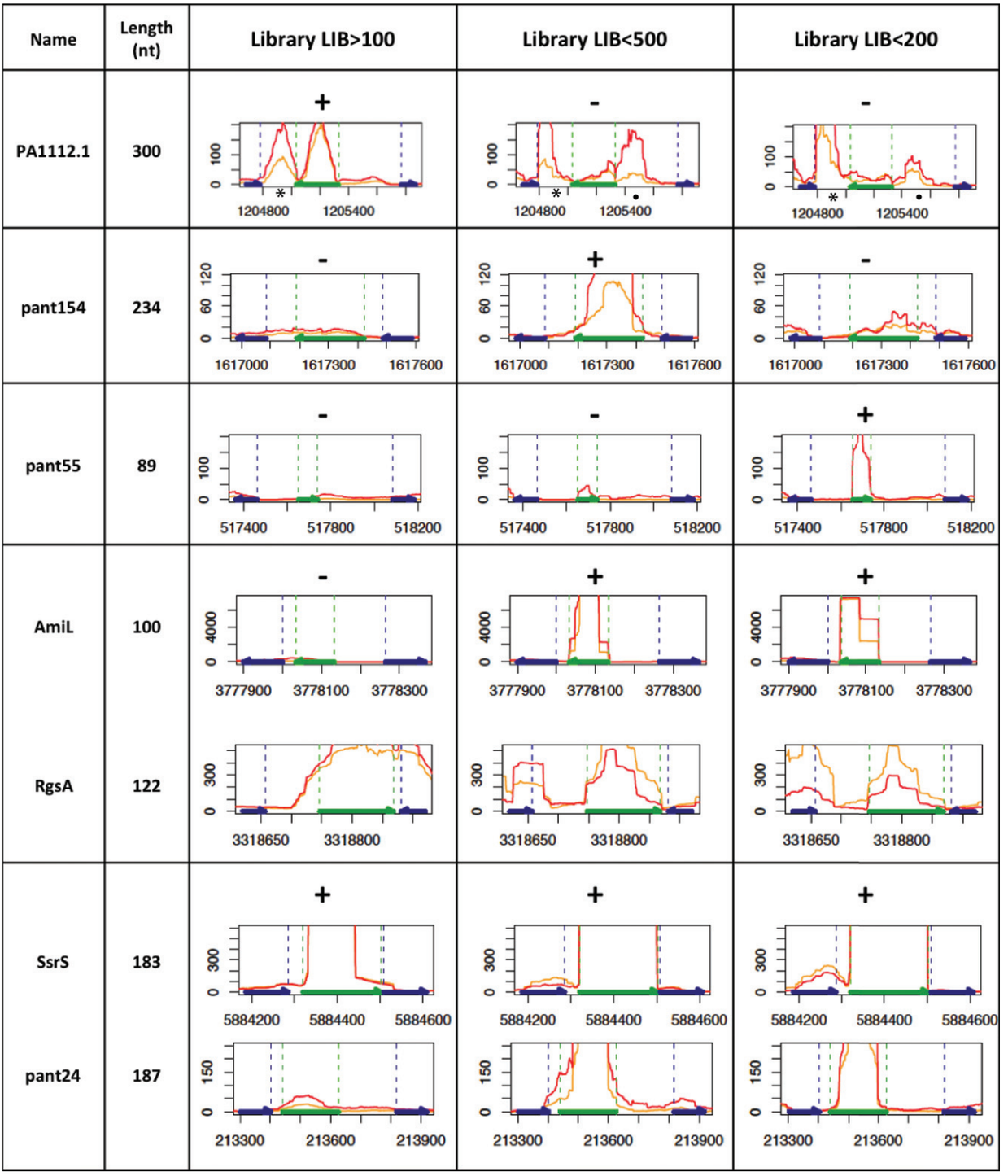


Fig. 4. Examples of sRNAs detected in the three different sequencing libraries LIB>100, LIB<500 and LIB<200. The *y*-axes of the plots denote normalized coverage at each nucleotide position. The *x*-axes denote the genomic positions according to the *P. aeruginosa* PAO1 genome coordinates (NC_002516.2). The left column of plots corresponds to the results from library LIB>100, the middle column to library LIB<500 and the right column to library LIB<200. PA1112.1, AmiL, SsrS and RgsA are already annotated sRNAs in *P. aeruginosa*. Legend: –, not detected in the library; +, detected in the library; orange lines, exponential phase; red lines, early-stationary phase; dashed green lines, sRNA coordinates; dashed blue lines, intergenic region coordinates; blue arrows, direction of flanking genes; green arrows, direction of the sRNA (if known). The peaks next to PA1112.1 correspond to the novel sRNAs pant120 (*) and pant121 (•) detected in this study.

relative abundance will go undetected in libraries containing more and larger transcripts. Likewise, AmiL is only identified in the LIB<200 and LIB<500 libraries due to its size (100 nt). The RgsA sRNA would have been misidentified as a UTR in the LIB>100 library, as the expression level of the sRNA is similar to the flanking gene (Fig. 4). This does not occur in the LIB<500 and LIB<200 libraries because the flanking gene is not included due to the size selection steps. Two examples of transcripts detected in all three libraries are SsrS and pant24 (Fig. 4). The intensity of pant24 is lower in the LIB>100 library than in the LIB<500 and LIB<200 libraries, which illustrates the increased sensitivity of the latter two libraries.

sRNA orientation and validation

The data from the LIB<500 and LIB<200 libraries were used to determine the orientation of the sRNAs whose coordinates and orientation were previously defined. The annotated orientations of all these sRNAs were confirmed with two exceptions, PhrY and RgsA. In the sequencing data the orientation of the PhrY sRNA is unclear and the RgsA sRNA appears to be encoded on the minus strand, contrary to its previous annotation (Livny *et al*., 2006; González *et al*., 2008; Sonnleitner *et al*., 2008). RgsA was previously predicted using the computational tool RNAz (Sonnleitner *et al*., 2008) and detected with Northern blots (Livny *et al*., 2006; González *et al*., 2008). Here, 5′-RACE was performed to determine the coding strand of RgsA. The 5′ ends of RgsA could be validated on both strands, suggesting that there is another sRNA encoded just opposite of RgsA (Table S5).

The strand specificity of the LIB<500 and LIB<200 libraries was utilized to determine the orientation of the 513 novel sRNAs identified in this study. Almost all of the sRNAs (506/513) were identified in the LIB<500 and/or LIB<200 libraries. It was possible to determine the orientation of 388 of the 506 sRNAs. For the others, assignment of orientation was precluded by the inability to determine the accurate limits of the transcripts.

The coding strand and the 5′ ends of six novel sRNAs (pant44, pant66, pant217, pant381, pant441, pant503) were determined using 5′-RACE. These sRNAs were chosen because their direction could be predicted from the LIB<200 and LIB<500 libraries and the average expression levels of four of them (pant44, pant66, pant217, pant381) were close to the cut-offs used for transcript detection. In all cases the coding strand detected by RACE was the same as the one predicted by the RNA-seq data. Moreover, the genomic coordinates are identical to or fall within 12 nt of the transcript 5′ end predicted by sequencing (Table S5).

The sRNAs pant43 and pant304 have similar coordinates and identical orientations to P2 and P17 sRNAs respectively. These sRNAs were predicted computationally in an earlier study, but could not be validated by Northern blotting (Livny *et al*., 2006). Here, pant43 (P2) was detected in the LIB<200 library and pant304 (P17) was detected in the LIB<500 and LIB<200 libraries, indicating that the RNA-seq libraries specific for sRNA detection (libraries LIB<500 and LIB<200) are more sensitive than Northern blots.

Growth phase-dependent expression of sRNAs

A Student's *t*-test was performed on the average expression of the transcripts to determine those with differential expression between the two conditions tested. Of the 513 novel transcripts, 178 are differentially expressed between the conditions tested (Table S4). Around 80% of these were significantly more expressed in early-stationary relative to exponential phase. Out of the 44 previously identified sRNAs, three (1059, P27 and PA5316.1) were found to be differentially expressed between the conditions tested (Table S3). Importantly, the three sequencing libraries were very consistent in the quantification of transcripts. Of the many sRNAs detected in more than one library type, most showed similar expression patterns in all libraries in which they were detected, or in other words, the libraries were congruent in whether the sRNAs were differentially expressed or not. This suggests that RNA-seq can be reliably used to quantify transcript abundance and concurs with other studies in which RNA-seq has been proven to be a very good method to quantify expression levels. Marioni and colleagues investigated the technical variance associated with Illumina sequencing and compared its ability to identify differentially expressed genes with Affymetrix arrays (Marioni *et al*., 2008). Illumina sequencing data were found to be highly reproducible, with relatively little technical variation and comparable to that from the arrays in enabling identification of differentially expressed genes (Marioni *et al*., 2008).

Conservation and homologue identification of the novel sRNAs

The sequence conservation of novel sRNAs in other bacteria was investigated using the BLASTN algorithm. The sequences of 12% (62/513) of the sRNAs are conserved in other bacteria, and 4% (22/513) are conserved in organisms outside the Pseudomonadaceae family (Table S6). The data show that 88% of the sRNAs are not conserved outside *P. aeruginosa*, indicating that the extent of sequence conservation between sRNAs in *P. aeruginosa* and other *Pseudomonas* spp. is limited. As *P. aeruginosa* is the only pseudomonad that is an opportunistic human pathogen, it is likely that some of

the specific sRNAs play a role in pathogenesis or confer an advantage inside the host. An example of an sRNA specific to *P. aeruginosa* is PrrH, a heme-regulated RNA encoded by the two identical sRNA genes *prrF1* and *prrF2*. These genes have a unique tandem organization in *P. aeruginosa* but are encoded at distal loci in the genomes of all other sequenced pseudomonads (Oglesby-Sherrouse and Vasil, 2010). There are no orthologous sequences in *P. aeruginosa* to previously characterized sRNAs in enterobacteria such as *E. coli* and *Salmonella*, suggesting that the sequence conservation among these sRNAs does not extend to the pseudomonads. The results illustrate the limitations in identifying sRNAs using bioinformatic prediction methods that are often based on sequence conservation. Moreover, they are consistent with the view that organisms have evolved species-specific regulatory networks involving sRNAs, and that these are necessary for niche adaptation, including host interactions in the case of pathogens.

Concluding remarks

This study is the first to use RNA-seq to discover novel sRNAs in *P. aeruginosa* and utilized three types of sequencing libraries to maximize the number of detected sRNAs. Over 500 novel sRNA transcripts are identified, increasing the number of known sRNAs in *P. aeruginosa* by a factor of 12 and revealing a large set of potential regulatory molecules in an already highly regulated species. Two libraries that do not include larger transcripts (LIB<500 and LIB<200) are clearly superior in detecting sRNAs, whereas the LIB>100 library contains information on the full transcriptome. Although RNA-seq enables the identification of novel transcripts, it is not well suited for accurately defining transcript length. As RNA-seq allows for detection of transcripts expressed at low levels, the sensitivity of existing validation methods is therefore a limitation for further characterization of these transcripts. Roughly one-third of the novel transcripts are differentially expressed between exponential and early-stationary growth phases, underscoring that RNA-seq is a sensitive method to monitor changes in gene expression. The limited extent of sequence conservation among pseudomonads highlights the challenges of sRNA bioinformatic prediction and emphasizes the importance of sRNAs at the species level in monitoring changes in environmental conditions. The approach described here can be used to characterize the sRNAome of any bacterium under highly diverse growth and stress conditions. In conclusion, the multitude of sRNAs in *P. aeruginosa* suggests the presence of a vast and complex regulatory network that will be highly important for our understanding of an organism with such a versatile life style that can successfully move from the environment into the human body as a pathogen.

Experimental procedures

RNA isolation

Single colonies of *P. aeruginosa* strain PAO1 were grown overnight in Luria–Bertani (LB) medium at 37°C. The cultures were diluted to a starting OD_{600} of 0.01 in 50 ml of LB media and grown at 37°C in separate 250 ml baffled Erlenmeyer flasks with shaking at 240 r.p.m. Cells were harvested from cultures grown to an OD_{600} of 0.6 (exponential phase) or 3.0 (early-stationary phase). Harvested cells were mixed immediately with 0.2 volumes of STOP solution (95% ethanol, 5% phenol) and pelleted by centrifugation. Total RNA was extracted with Trizol (Invitrogen). Removal of DNA was carried out by treatment with DNase I (Fermentas) in combination with the RNase inhibitor RiboLock (Fermentas). The integrity of total RNA, the presence of 5S rRNA and tRNAs, and DNA contamination were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

Removal of 23S, 16S and 5S rRNAs

The 23S, 16S and 5S rRNAs were removed by subtractive hybridization using the MICROBExpress kit (Ambion) with modifications. Capture oligonucleotides complementary to the rRNAs were designed specifically for *P. aeruginosa* (Table S2). For preparation of library LIB>100 (described below), 5 μM of each capture oligonucleotide was used, for a total capture oligonucleotide concentration of 25 μM. Compared with the standard protocol, 25% more capture oligonucleotides and magnetic beads were used. For preparation of libraries LIB<500 and LIB<200 (described below), 5 μM of the 5S rRNA capture oligonucleotide and 50% less magnetic beads were used relative to the standard protocol. Removal of rRNAs was confirmed with an Agilent 2100 Bioanalyzer (Agilent Technologies) (Fig. S1).

Library preparation and RNA sequencing

The sequencing libraries LIB>100, LIB<500 and LIB<200 were constructed using three different protocols (Fig. 1, Table 1). Each library type was prepared with RNA isolated from cells grown in triplicate for each condition (exponential and early-stationary phase). After each step the samples were validated using an Agilent 2100 Bioanalyzer (Agilent Technologies), and the final concentration was measured using a Qubit 2.0 Fluorometer (Invitrogen). The libraries were sequenced using the Illumina HiSeq2000 platform with a paired-end protocol and read lengths of 100 nt.

LIB>100 library. This type of library contains transcripts longer than 100 nt and was prepared using the TruSeq RNA Sample Preparation kit (Illumina). Briefly, the rRNA-depleted RNA is fragmented using divalent cations under elevated temperature. The cleaved RNA fragments are copied into

cDNA using reverse transcriptase and random primers, followed by second-strand cDNA synthesis using DNA polymerase I and RNase H. After this step, transcripts shorter than 100 nt are not retained due to the purification performed with the Agencourt AMPure XP beads (Beckman Coulter Genomics). The cDNA fragments then go through an end repair process, 3′-addition of single 'A' bases and adapter ligation. This is followed by product purification and PCR amplification to generate the final cDNA library.

LIB<500 library. This type of library contains transcripts shorter than 500 nt. RNA size selection was performed by running total RNA on 10% polyacrylamide gels containing 10 M urea. Gel slices corresponding to RNAs up to 500 nt were excised, followed by elution of RNA in 0.4 M NaCl and precipitation with ethanol. The 5S rRNA was depleted as previously described, followed by treatment with Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37°C for 90 min. Tobacco Acid Pyrophosphatase (TAP) is used to convert 5′-triphosphate RNA into 5′-monophosphate RNA, which is important for correct adapter ligation. This was followed by treatment with RNase III (Ambion) for 10 min at 37°C to fragment the RNA. RNase III fragments RNA into smaller pieces containing a 5′-phosphoryl group and a 3′-hydroxyl group, which is important for specific adapter ligation in the next step. Sequential ligation of RNA 3′ and 5′ adapters was performed using the adapters and enzymes from the TruSeq Small RNA Sample Preparation kit (Illumina). Next, reverse transcription followed by PCR amplification was performed to form cDNA constructs based on the RNA fragments ligated with 3′ and 5′ adapters, selectively enriching fragments with adapter molecules on both ends. The reverse transcription and subsequent PCR amplification were performed using the enzymes and reagents from the TruSeq Small RNA Sample Preparation kit (Illumina). Agencourt AMPure XP beads (Beckman Coulter Genomics) were used for the post-PCR clean-up. In this and the LIB<200 libraries, the fragments corresponding to transcripts shorter than 100 nt are retained due to ligated RNA adapters that increase the length of the fragments by 125 nt.

LIB<200 library. This type of library contains transcripts shorter than 200 nt. Size selection of RNAs up to 200 nt from total RNA was carried out with the mirVana miRNA Isolation Kit (Ambion), followed by depletion of 5S rRNA as previously described. The TAP treatment, RNA adapter ligation, reverse transcription and PCR amplification and clean-up steps were performed as described for library LIB<500.

Read mapping

Reads were mapped onto the *P. aeruginosa* PAO1 genome (RefSeq Accession No. NC_002516.2) using the Bowtie short read aligner (Langmead *et al*., 2009). For the LIB<500 and LIB<200 libraries, the adapters had to be trimmed in the reads with inserts shorter than 100 nt. Reads with only one alignment reported were retained. Read alignments from Bowtie were handled using SAMtools (Li *et al*., 2009). Read alignments onto the reference genome were visualized using Tablet 1.11.08.10 (Milne *et al*., 2010).

Transcript quantification

In order to obtain normalized expression intensities of the read coverage depth at each position in the genome, the number of reads in each replicate was normalized according to the total number of reads in the library, and expression intensities from replicate samples were averaged. The average expression of each already annotated sRNA transcript was calculated by dividing the normalized coverage depth of the transcript by the length of the transcript.

Identification of intergenic sRNAs

Already annotated sRNAs were identified by visual inspection of the read coverage of the reference genome at the relevant loci. A custom-made script was designed to detect novel transcripts in intergenic regions. According to this, a transcript was detected if the coverage depth at each base was higher than a certain cut-off. The cut-off was set for each library based on the average expression intensities of the 44 previously validated sRNAs in *P. aeruginosa* in each library type (Table S7). In addition, the average expression intensities of the detected transcript had to be twice the previously mentioned cut-off. This ensures detection of abrupt increases in read coverage depths. To define the length of the transcripts more accurately, the lowest chosen cut-off (the one for library LIB>100) was applied to the peaks previously found in libraries LIB<500 and LIB<200. To prevent misclassification of UTRs as ncRNAs, the script only detects transcripts that begin and end within an intergenic region. Transcripts of at least 100 nt (LIB>100) or 70 nt (LIB<500, LIB<200) were considered further. The resulting transcripts from automatic classification were re-evaluated by visual inspection using Tablet 1.11.08.10 (Milne *et al*., 2010).

Differential expression of intergenic sRNAs

The average expression of each sRNA was calculated by dividing the normalized read coverage depth by the length of the transcript. A two-sample Student's *t*-test was performed on the average expression of the sRNAs to determine those with differential expression between the two conditions tested (P-value < 0.05 and fold change > 2). The Benjamini–Hochberg multiple testing correction was applied (Benjamini, 1995). The triplicates of the conditions were not averaged for this analysis. The expression of a sRNA was considered significantly different, if the sRNA was differentially expressed in all the libraries in which it was detected.

Homologue identification

The novel sRNAs were searched against the Rfam database to look for homologies to ncRNA entries. Putative homologues of each novel transcriptional unit were predicted via BLASTN comparison of each sequence to all sequenced bacterial genomes ($E = 10^{-6}$, word = 11). Only hits with nucleotide identity higher than 60% combined with coverage between query and subject sequence higher than 80% were considered to be conserved.

Validation of coding strand and 5′ ends of sRNAs

We applied a RACE procedure to determine the 5′ ends of a selection of the novel sRNAs. The protocol used is based on the method of Gerhart and colleagues with modifications (Gerhart *et al*., 2005). In brief, DNase-treated total RNA (described above) was incubated with Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37°C for 60 min. A short RNA adapter was ligated to the 5′ ends using T4 RNA ligase (New England Biolabs). The 5′ adapter-ligated RNA was reverse-transcribed using a primer complementary to the sRNA (GSP1) and the Thermoscript RT-PCR system (Invitrogen). Primers were removed using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The resulting cDNA was used as the template in PCR reactions using strand-specific primers to the sRNA (GSP2) along with an adapter-specific primer. Negative controls were performed using the adapter-ligated RNA as template. The oligonucleotides used for RACE are listed in Table S2. PCR clean-up was carried out using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The PCR products were resolved using an Agilent 2100 Bioanalyzer High Sensitivity DNA assay (Agilent Technologies) to determine their lengths. Products were sequenced with the primer used in PCR (GSP2) and with the adapter-specific primer by LGC Genomics GmbH (Germany).

Acknowledgements

Björn Renberg of the Royal Institute of Technology (KTH) and the Science for Life Laboratory in Stockholm are gratefully acknowledged for helpful discussions and sequencing of two sets of libraries. Anna Koza is thanked for assistance with sRNA validation. Laurent Gautier and Khoa Nguyen Do of the DTU Multi-Assay Core are thanked for their help and use of their facilities in connection with library preparation. This work was supported by the Novo Nordisk Foundation, Lundbeck Foundation and Danish National Research Agency.

References

Albrecht, M., Sharma, C.M., Reinhardt, R., Vogel, J., and Rudel, T. (2010) Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res* **38:** 868–877.

Bastet, L., Dubé, A., Massé, E., and Lafontaine, D.A. (2011) New insights into riboswitch regulation mechanisms. *Mol Microbiol* **80:** 1148–1154.

Benjamini, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B* **57:** 289–300.

Bradley, E.S., Bodi, K., Ismail, A.M., and Camilli, A. (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in *Vibrio cholerae*. *PLoS Pathog* **7:** e1002126.

Croucher, N.J., and Thomson, N.R. (2010) Studying bacterial transcriptomes using RNA-seq. *Curr Opin Microbiol* **13:** 619–624.

Gerhart, E., Wagner, H., and Vogel, J. (2005) Approaches to identify novel non-messenger RNAs in bacteria and to investigate their biological functions: functional analysis of identified non-mRNAs. In *Handbook of RNA Biochemistry*. Hartmann, R.K., Bindereif, A., Schön, A., and Westhof, E. (eds). Weinheim: Wiley-VCH Verlag GmbH, pp. 614–642.

González, N., Heeb, S., Valverde, C., Kay, E., Reimmann, C., Junier, T., and Haas, D. (2008) Genome-wide search reveals a novel GacA-regulated small RNA in *Pseudomonas* species. *BMC Genomics* **9:** 167.

Gottesman, S., and Storz, G. (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* **3:** a00379.

He, S., Wurtzel, O., Singh, K., Froula, J.L., Yilmaz, S., Tringe, S.G., *et al*. (2010) Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Methods* **7:** 807–812.

Irnov, I., Sharma, C.M., Vogel, J., and Winkler, W.C. (2010) Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Res* **38:** 6637–6651.

Karpinets, T., Greenwood, D., Sams, C., and Ammons, J. (2006) RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC Biol* **4:** 30.

Koo, J.T., Alleyne, T.M., Schiano, C.A., Jafari, N., and Lathem, W.W. (2011) Global discovery of small RNAs in *Yersinia pseudotuberculosis* identifies *Yersinia*-specific small, noncoding RNAs required for virulence. *Proc Natl Acad Sci USA* **108:** E709–E717.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., *et al*. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25:** 2078–2079.

Liu, J.M., Livny, J., Lawrence, M.S., Kimball, M.D., Waldor, M.K., and Camilli, A. (2009) Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic Acids Res* **37:** e46.

Livny, J., Brencic, A., Lory, S., and Waldor, M.K. (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Res* **34:** 3484.

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., *et al*. (2009) A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* **139:** 770–779.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* **18:** 1509–1517.

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., and Marshall, D. (2010) Tablet – next generation sequence assembly visualization. *Bioinformatics* **26:** 401–402.

Mitschke, J., Georg, J., Scholz, I., Sharma, C.M., Dienst, D., Bantscheff, J., *et al*. (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci USA* **108:** 2124–2129.

Mraheil, M.A., Billion, A., Mohamed, W., Mukherjee, K., Kuenne, C., Pischimarov, J., *et al*. (2011) The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res* **39:** 4235–4248.

Oglesby-Sherrouse, A.G., and Vasil, M.L. (2010) Characterization of a heme-regulated non-coding RNA encoded by the *prrF* locus of *Pseudomonas aeruginosa*. *PLoS ONE* **5:** e9930.

Oliver, H.F., Orsi, R.H., Ponnala, L., Keich, U., Wang, W., Sun, Q., *et al*. (2009) Deep RNA sequencing of *L. monocytogenes* reveals overlapping and extensive stationary phase and sigma B-dependent transcriptomes, including multiple highly transcribed noncoding RNAs. *BMC Genomics* **10:** 641.

Papenfort, K., and Vogel, J. (2010) Regulatory RNA in bacterial pathogens. *Cell Host Microbe* **8:** 116–127.

Perkins, T.T., Kingsley, R.A., Fookes, M.C., Gardner, P.P., James, K.D., Yu, L., *et al*. (2009) A strand-specific RNA–Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. *PLoS Genet* **5:** e1000569.

Repoila, F., and Darfeuille, F. (2009) Small regulatory noncoding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell* **101:** 117–131.

Rudinger-Thirion, J., Lescure, A., Paulus, C., and Frugier, M. (2011) Misfolded human tRNA isodecoder binds and neutralizes a 3′ UTR-embedded Alu element. *Proc Natl Acad Sci USA* **108:** 794–802.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., *et al*. (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464:** 250–255.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., *et al*. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4:** e1000163.

Sonnleitner, E., and Haas, D. (2011) Small RNAs as regulators of primary and secondary metabolism in *Pseudomonas* species. *Appl Microbiol Biotechnol* **91:** 63–79.

Sonnleitner, E., Sorger-Domenigg, T., Madej, M.J., Findeiss, S., Hackermüller, J., Hüttenhofer, A., *et al*. (2008) Detection of small RNAs in *Pseudomonas aeruginosa* by RNomics and structure-based bioinformatic tools. *Microbiology* **154:** 3175–3187.

Sonnleitner, E., González, N., and Haas, D. (2010) Small RNAs of *Pseudomonas* spp. In *Pseudomonas – Molecular Microbiology, Infection and Biodiversity*, Vol. 6. Ramos, J.L., and Filloux, A. (eds). Dordrecht: Springer Netherlands, pp. 3–28.

Storz, G., Vogel, J., and Wassarman, K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43:** 880–891.

Toschka, H.Y., Struck, J.C., and Erdmann, V.A. (1989) The 4.5S RNA gene from *Pseudomonas aeruginosa*. *Nucleic Acids Res* **17:** 31–36.

Vockenhuber, M., Sharma, C.M., Statt, M.G., Schmidt, D., Xu, Z., Dietrich, S., *et al*. (2011) Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*. *RNA Biol* **8:** 468–477.

Vogel, D.W., Hartmann, R.K., Struck, J.C.R., Ulbrich, N., and Erdmann, V.A. (1987) The sequence of the 6S RNA gene of *Pseudomonas aeruginosa*. *Nucleic Acids Res* **15:** 4583–4591.

Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136:** 615–628.

Weissenmayer, B.A., Prendergast, J.G.D., Lohan, A.J., and Loftus, B.J. (2011) Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs. *PLoS ONE* **6:** e17570.

Wilderman, P.J., Sowa, N.A., FitzGerald, D.J., FitzGerald, P.C., Gottesman, S., Ochsner, U.A., and Vasil, M.L. (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc Natl Acad Sci USA* **101:** 9792–9797.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Agilent Bioanalyzer analysis of total and rRNA-depleted RNA. The 23S, 16S and 5S rRNAs were removed from the total RNA (A) by subtractive hybridization as described in the *Experimental procedures*. Note that the peaks corresponding to 23S, 16S and 5S rRNAs are not present in the rRNA-depleted sample (B). The peaks marked with asterisks (*) probably correspond to tRNAs and sRNA species shorter than 120 nt.

Table S1. Sequencing results.

Table S2. Oligonucleotides used in this study.

Table S3. Annotated sRNAs in *Pseudomonas aeruginosa*.

Table S4. Intergenic sRNAs candidates in *Pseudomonas aeruginosa*.

Table S5. 5′-RACE validation of sRNA candidates.

Table S6. Novel sRNA candidates conserved in organisms outside the Pseudomonadaceae family.

Table S7. Cut-offs used for the detection of novel transcripts.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Supplementary Table 1. Sequencing results.

Library	Condition	Number of replicates	Average total number of reads	Average number of mapped reads	Average % of read coverage corresponding to			
					23S rRNA	16S rRNA	5S rRNA	tRNAs
LIB>100	Exponential	3	16961351	14676650	12.60%	6.74%	0.07%	2.53%
	Early-stationary	3	24030975	21001129	15.60%	7.88%	0.01%	1.15%
LIB<500	Exponential	3	16577877	13119470	3.50%	1.62%	0.16%	3.69%
	Early-stationary	3	14758047	10822645	1.79%	1.29%	0.19%	3.28%
LIB<200	Exponential	3	12822325	7564110	5.18%	3.01%	0.25%	3.80%
	Early-stationary	3	15070347	6755123	3.91%	2.62%	0.21%	1.89%
		Average	**16703487**	**12323187**	**7.10%**	**3.86%**	**0.15%**	**2.72%**

Supplementary Table 2. Oligonucleotides used in this study.

Name	Sequence (5' -> 3')	Use
23S-1954	AAAAAAAAAAAAAAAAAAACTTACCCGACAAGGAATTTCGC	Removal of 23S rRNA
23S-2511	AAAAAAAAAAAAAAAAAAGAGCCGACATCGAGGTGCCAAAC	Removal of 23S rRNA
16S-807	AAAAAAAAAAAAAAAAAATGGACTACCAGGGTATCTAATCC	Removal of 16S rRNA
16S-1114	AAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTACGGGACTT	Removal of 16S rRNA
5S	AAAAAAAAAAAAAAAAAAGCGTTTCACTTCTGAGTTCGGCA	Removal of 5S rRNA
RNA adapter	AUAUGCGCGAAUUCCUGUAGAACGAACACUAGAAGAAA	RNA ligation in 5'-RACE
Adapter-specific primer	GCGCGAATTCCTGTAGA	PCR in 5'-RACE
GSP1 pant44	TCAGGAAGCGTTCTTGTTGG	RT in 5'-RACE
GSP2 pant44	AGCGTTCTTGTTGGATTTCTAGC	PCR in 5'-RACE
GSP1 pant66	GCGAGAATCCCCCAGACT	RT in 5'-RACE
GSP2 pant66	GAATCCCCCAGACTCCGT	PCR in 5'-RACE
GSP1 pant217	GTACGGCAACCCCTGCAT	RT in 5'-RACE
GSP2 pant217	GCAACCCCTGCATCGACCG	PCR in 5'-RACE
GSP1 pant381	AGGAGACAGGTTGGGA	RT in 5'-RACE
GSP2 pant381	GACAGGTTGGGAGGGAGTT	PCR in 5'-RACE
GSP1 pant441	GCGACTCCCCAAGCATAGCCG	RT in 5'-RACE
GSP2 pant441	CCAAGCATAGCCGGCGGGAC	PCR in 5'-RACE
GSP1 pant503	CTGGCACTACGACGAATGG	RT in 5'-RACE
GSP2 pant503	CACTACGACGAATGGCCAAC	PCR in 5'-RACE
GSP1 RgsA (forward)	ATCCGGGTCAAGACCAT	RT in 5'-RACE
GSP2 RgsA (forward)	GTCAAGACCATTAGGAG	PCR in 5'-RACE
GSP1 RgsA (reverse)	AGTACCCAGCCAGTGAAG	RT in 5'-RACE
GSP2 RgsA (reverse)	CAGCCAGTGAAGGTAC	PCR in 5'-RACE

Supplementary Table 3. Annotated sRNAs in *Pseudomonas aeruginosa*.

Name	Start[a]	End[a]	Intergenic region start	Intergenic region end	Length[b]	5' flanking gene	3' flanking gene	Direction[c]	Distance to 5' flanking gene (nt)	Distance to 3' flanking gene (nt)	Detection in Library A	Detection in Library B	Detection in Library C	Differential expression[d]	Fold change[e]	Reference
P1	-	-	334456	334733	300	PA0296	PA0297	<<>	?	?	-	+	+	-	-	(Livny et al., 2006)
rsmY	586867	586990	586664	587016	124	PA0527	PA0528	<><	203	26	-	+	+	-	-	(Kay et al., 2006; Livny et al., 2006; Sonnleitner et al., 2008)
phrD	785498	785570	785174	785969	73	PA0714	PA0715	>>>	324	399	-	+	+	-	-	(Sonnleitner et al., 2008)
ssrA	901520	901872	901047	901933	353[f]	PA0826	PA0827	<<<	473	61	+	+	+	-	-	(Sonnleitner et al., 2008; Wiilliams and Bartel 1996)
P5	-	-	912780	913085	90	PA0836	PA0837	<>>	?	?	+	+	+	-	-	(González et al., 2008; Livny et al., 2006)
P7	-	-	971625	972166	140	PA0887	PA0888	>>>	?	?	-	+	+	-	-	(Livny et al., 2006)
P8	-	-	1117319	1118157	130	PA1030	PA1031	>>>	?	?	+	+	+	-	-	(Livny et al., 2006)
sRNA622	-	-	1140860	1141267	200,300	PA1052	PA1053	>>>	?	?	+	+	+	-	-	(González et al., 2008)
PA1112.1	1205031	1205330	1204782	1205770	300	PA1112	PA1113	><>	249	440	+	-	-	-	-	(González et al., 2008)
P9	1436491	1436618	1436397	1436663	128	PA1324	PA1325	>>>	94	45	+	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
ffs	1668935	1669048	1668833	1669085	173 (113)[g]	PA1530	PA1531	>><	102	37	-	+	+	-	-	(Toschka 1989; González et al., 2008; Sonnleitner et al., 2008)
P11	-	-	1928627	1928893	100	PA1781	PA1782	<<<	?	?	-	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
sRNA1059	-	-	1996807	1997508	150,200,300	PA1838	PA1839	<<<	?	?	+	+	+	+	9	(González et al., 2008)
sRNA1466	-	-	2918212	2918603	300	PA2581	PA2581.1	>?<	?	?	+	-	-	-	-	(González et al., 2008; Sonnleitner et al., 2008)
PA2744.1	3106919	3106994	3106752	3107002	76	PA2744	PA2745	<<<	1	60	-	-	-	-	-	(Livny et al., 2006)
sRNA1559	-	-	3112151	3112876	250,300	PA2750	PA2751	>?>	?	?	-	+	+	-	-	(González et al., 2008)
P14	-	-	3206497	3206915	300	PA2852.1	PA2853	<>>	?	?	-	+	+	-	-	(Livny et al., 2006)
P15	-	-	3298922	3299492	180	PA2942	PA2943	<>>	?	?	+	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
rgsA	3318747	3318868	3318657	3318881	122	PA2958	PA2959	>><	90	13	-	+	+	-	-	(Livny et al., 2006; González et al., 2008; Sonnleitner et al., 2008)
sRNA1714	-	-	3360654	3360873	200	PA3001	PA3002	<?>	?	?	-	-	+	-	-	(González et al., 2008)
P18	-	-	3702951	3703166	100	PA3304	PA3305	<<<	?	?	-	-	+	-	-	(Livny et al., 2006)
phrS	3705309	3705521	3705161	3705888	213	PA3305	PA3306	<<<	148	367	+	+	+	-	-	(Livny et al., 2006; González et al., 2008; Sonnleitner et al., 2008)
amiL	3778034	3778133	3778000	3778265	100	PA3366	PA3367	<<>	34	132	-	+	+	-	-	(Sonnleitner et al., 2008)
rsmZ	4057543	4057658	4057483	4057910	116	PA3621	PA3622	>><	60	252	-	+	+	-	-	(Heurlier et al., 2004; Livny et al., 2006)
P24	-	-	4444597	4444977	300	PA3964	PA3965	><>	?	?	+	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
sRNA2315	-	-	4536493	4536919	180	PA4055	PA4056	<<<	?	?	+	+	+	-	-	(González et al., 2008; Sonnleitner et al., 2008)
P26	-	-	4780618	4780838	250	PA4270	PA4271	<<<	?	?	-	-	-	-	-	(Livny et al., 2006)
P27	-	-	4781786	4781985	90	PA4272	PA4273	<><	?	?	-	+	-	+	0.3	(Livny et al., 2006)
PA4406.1	4939194	4939277	4939188	4939299	84	PA4406	PA4407	<><	6	22	-	-	+	-	-	(Sonnleitner et al., 2008)
rnpB	-	-	4956029	4956732	350	PA4421	PA4422	<<<	?	?	+	+	+	-	-	(Livny et al., 2006; González et al., 2008; Sonnleitner et al., 2008)
PA4451.1	4985782	4985843	4985731	4985846	62	PA4451	PA4452	<<<	51	3	-	-	-	-	-	(Livny et al., 2006)
sRNA2626	-	-	5196833	5197184	200	PA4628	PA4629	>>>	?	?	+	+	+	-	-	(González et al., 2008; Sonnleitner et al., 2008)
prrH	5283960	5284319	5283906	5284368	360	PA4703	PA4704	>><	54	49	-	+	+	-	-	(Oglesby-Sherrouse and Vasil et al., 2010)
prrF1	5283960	5284110	5283906	5284368	151	PA4704	PA4705	>><	54	258	-	+	+	-	-	(Livny et al., 2006; González et al., 2008; Sonnleitner et al., 2008; Wilderman et al., 2004)
prrF2	5284172	5284319	5283906	5284368	148	PA4704	PA4705	>><	266	49	-	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
crcZ	5308587	5308993	5308425	5309326	407	PA4727	PA4728	>>>	162	333	+	+	+	-	-	(Sonnleitner et al., 2009)
P30	-	-	5308425	5309326	180	PA4726	PA4727	><>	?	?	+	+	+	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
P32	-	-	5344904	5345085	80	PA4758	PA4759	<<<	?	?	-	-	-	-	-	(Livny et al., 2006; Sonnleitner et al., 2008)
P34	-	-	5835071	5835481	150	PA5181	PA5182	><>	?	?	-	+	+	-	-	(Livny et al., 2006)
PhrX	5836429	5836579	5836402	5836909	151	PA5183	PA5184	><>	27	330	+	+	+	-	-	(Sonnleitner et al., 2008)
PhrY	5859480	5859674	5859457	5859792	195	PA5204	PA5205	<>>	23	118	+	+	+	-	-	(Sonnleitner et al., 2008)
ssrS	5884320	5884502	5884286	5884508	183	PA5227	PA5228	>>>	34	6	+	+	+	-	-	(Sonnleitner et al., 2008; Vogel et al., 1987)
PA5316.1	5986120	5986170	5986120	5986474	51	PA5316	PA5317	<>>	0	304	-	-	-	+	0.3	(Sonnleitner et al., 2008)
Putative spot42	-	-	6183520	6183784	265	PA5492	PA5493	>><	0	0	+	+	+	-	-	(Gottesman et al., 2006)

a A location is given when the 5' and 3' ends can be deduced with reasonable certainty from available data.

b Experimentally found values.

c The middle arrow indicates the orientation of the sRNA, while the flanking arrows indicate the orientation of the adjacent genes. Unknown orientation of sRNAs is indicated by a question mark.

d A + indicates that the sRNA is differentially expressed in the conditions tested in this work.

e The fold change is calculated as the ratio of the intensities of the sRNA between the early-stationary and exponential conditions. Thus, a fold change > 1 means that the sRNA is more expressed in the early-stationary than the exponential phase.

f Mature tmRNA according to sequence homologies (Williams and Bartel, 1996).

g The mature form was reported with a length of 113 nt (Toschka et al., 1989).

REFERENCES

González, N., Heeb, S., Valverde, C., Kay, E., Reimmann, C., Junier, T., and Haas, D. (2008) Genome-wide search reveals a novel GacA-regulated small RNA in Pseudomonas species. *BMC genomics* **9**: 167.

Gottesman, S., McCullen, C.A., Guillier, M., Vanderpool, C.K., Majdalani, N., Benhammou, J. et al. (2006) Small RNA regulators and the bacterial response to stress. *Cold Spring Harbor symposia on quantitative biology* **71**: 1.

Heurlier, K., Williams, F., Heeb, S., Dormond, C., Pessi, G., Singer, D. et al. (2004) Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in Pseudomonas aeruginosa PAO1. *Journal of Bacteriology* **186**: 2936-2945.

Kay, E., Humair, B., Dénervaud, V., Riedel, K., Spahr, S., Eberl, L. et al. (2006) Two GacA-Dependent Small RNAs Modulate the Quorum-Sensing Response in Pseudomonas aeruginosa. *Journal of Bacteriology* **188**: 6026-6033.

Livny, J., Brencic, A., Lory, S., and Waldor, M.K. (2006) Identification of 17 Pseudomonas aeruginosa sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Research* **34**: 3484.

Oglesby-Sherrouse, A.G., and Vasil, M.L. (2010) Characterization of a heme-regulated non-coding RNA encoded by the prrF locus of Pseudomonas aeruginosa. *PLoS ONE* **5**: e9930.

Sonnleitner, E., Sorger-Domenigg, T., Madej, M.J., Findeiss, S., Hackermüller, J., Hüttenhofer, A. et al. (2008) Detection of small RNAs in Pseudomonas aeruginosa by RNomics and structure-based bioinformatic tools. *Microbiology* **154**: 3175-3187.

Sonnleitner, E., Abdou, L., and Haas, D. (2009) Small RNA as global regulator of carbon catabolite repression in Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences* **106**: 21866-21871.

Toschka, H.Y., Struck, J.C., and Erdmann, V.A. (1989) The 4.5S RNA gene from Pseudomonas aeruginosa. *Nucleic Acids Research* **17**: 31-36.

Vogel, D.W., Hartmann, R.K., Struck, J.C.R., Ulbrich, N., and Erdmann, V.A. (1987) The sequence of the 6S RNA gene of Pseudomonas aeruginosa. *Nucleic Acids Research* **15**: 4583-4591.

Wilderman, P.J., Sowa, N.A., FitzGerald, D.J., FitzGerald, P.C., Gottesman, S., Ochsner, U.A., and Vasil, M.L. (2004) Identification of tandem duplicate regulatory small RNAs in P. aeruginosa involved in iron homeostasis. *Proceedings of the National Academy of Sciences of the USA* **101**: 9792-9797.

Williams, K.P., and Bartel, D.P. (1996) Phylogenetic analysis of tmRNA secondary structure. *RNA* **2**: 1306-1310.

Supplementary Table 4. Intergenic sRNAs candidates in *Pseudomonas aeruginosa*.

Name	Start[a]	End[a]	Intergenic region start	Intergenic region end	Length[b]	5' flanking gene	3' flanking gene	Direction[c]	Distance to 5' flanking gene (nt)	Distance to 3' flanking gene (nt)	Detection in Library A	Detection in Library B	Detection in Library C	Differential expression[d]	Fold change[e]	Rfam[f]	blastn[g]
pant1	46	293	1	482	248	PA5570	PA0001	<>>	45	189	+	+	-	-	-	-	-
pant2	8453	8654	8340	8670	202	PA0006	PA0007	<>>	113	16	-	+	+	+	4.9	-	-
pant3	40967	41110	40817	41112	144	PA0039	PA0040	>>>	150	2	-	-	+	-	-	-	-
pant4	53658	54002	53522	56545	345	PA0041	PA0042	>?>	136	2543	-	+	+	-	-	-	-
pant5	54656	54806	53522	56545	151	PA0041	PA0042	>>>	1134	1739	-	+	-	+	8.5	-	-
pant6	55066	56182	53522	56545	1117	PA0041	PA0042	>?>	1544	363	+	+	-	-	-	-	-
pant7	60222	60564	60160	60655	343	PA0044	PA0045	>?>	62	91	-	+	+	-	-	-	-
pant8	62798	63023	62787	63067	226	PA0048	PA0049	<<<	11	44	-	-	+	+	13.2	-	-
pant9	64880	65034	64730	65338	155	PA0049	PA0050	<<>	150	304	-	+	+	+	10.3	-	-
pant10	65587	66281	65480	66302	695	PA0050	PA0051	>?>	107	21	+	+	+	-	-	-	-
pant11	68822	68972	68617	69271	151	PA0052	PA0053	<?>	205	299	-	-	+	-	-	-	-
pant12	123728	123839	123496	123870	112	PA0101	PA0102	>?>	232	31	-	-	+	+	4.3	-	-
pant13	124638	124776	124600	124809	139	PA0102	PA0103	>>>	38	33	-	+	+	+	0.3	-	-
pant14	127119	127375	127115	127377	257	PA0104	PA0105	>>>	4	2	-	+	+	+	11.9	-	-
pant15	143369	143529	143268	143566	161	PA0123	PA0124	<?<	101	37	-	+	+	-	-	-	-
pant16	153324	153524	153307	153695	201	PA0134	PA0135	<?>	17	171	-	+	+	-	-	-	-
pant17	161581	161712	161449	161905	132	PA0141	PA0142	<<>	132	193	-	+	+	-	-	-	-
pant18	176133	176214	176109	176314	82	PA0154	PA0155	><>	24	100	-	+	-	+	2.7	-	-
pant19	182619	182692	182570	182767	74	PA0158	PA0159	><>	49	75	-	-	+	-	-	-	-
pant20	188728	188861	188449	189119	134	PA0164	PA0165	>>>	279	258	-	+	+	+	0.3	-	-
pant21	188903	189044	188449	189119	142	PA0164	PA0165	>>>	454	75	-	+	-	+	0.4	-	-
pant22	193682	193797	193673	193798	116	PA0169	PA0170	<<<	9	1	-	+	-	-	-	-	-
pant23	196750	196823	196749	197010	74	PA0172	PA0173	<><	1	187	-	+	-	+	0.5	-	-
pant24	213440	213626	213402	213818	187	PA0186	PA0187	>?>	38	192	+	+	+	-	-	-	-
pant25	218894	218999	218823	219171	106	PA0191	PA0192	<>>	71	172	-	+	-	-	-	-	-
pant26	223577	223723	223455	224100	147	PA0194	PA0195	>?>	122	377	-	+	+	-	-	-	-
pant27	223784	224041	223455	224100	258	PA0194	PA0195	>?>	329	59	-	+	+	+	6.7	-	-
pant28	245760	245915	245620	245947	156	PA0217	PA0218	<?<	140	32	+	+	+	-	-	-	-
pant29	264600	264954	264523	264963	355	PA0234	PA0235	>><	77	9	-	+	+	-	-	-	+
pant30	269579	269667	269520	269668	89	PA0238	PA0239	>?>	59	1	-	+	+	-	-	-	-
pant31	288131	288320	288121	288383	190	PA0256	PA0257	<><	10	63	-	+	+	-	-	-	-
pant32	291005	291139	291005	291153	135	PA0259	PA0260	<<<	0	14	-	+	+	-	-	-	-
pant33	297591	297937	297562	298815	347	PA0263	PA0263.1	<><	29	878	-	+	+	-	-	-	-
pant34	298455	298687	297562	298815	233	PA0263	PA0263.1	<?<	893	128	+	-	-	-	-	-	-
pant35	301059	301154	300974	301217	96	PA0265	PA0266	>>>	85	63	-	+	-	-	-	gabT	-
pant36	317834	317961	317790	317965	128	PA0283	PA0284	<?<	44	4	-	-	+	-	-	-	-
pant37	326764	327071	326672	327283	308	PA0290	PA0291	>>>	92	212	-	+	+	-	-	-	-
pant38	327152	327252	326672	327283	101	PA0290	PA0291	>?>	480	31	-	+	-	-	-	-	-
pant39	339649	339764	339541	339958	116	PA0300	PA0301	>>>	108	194	-	+	+	+	3.8	-	-
pant40	347931	348110	347854	349049	180	PA0306	PA0308	><>	77	939	-	+	+	-	-	-	-
pant41	348843	349035	347854	349049	193	PA0306	PA0308	><>	989	14	+	+	+	-	-	-	+
pant42	354489	354679	354462	354753	191	PA0314	PA0315	<?>	27	74	-	-	+	+	2.6	-	++
pant43	356490	356599	356478	356680	110	PA0316	PA0317	<?>	12	81	-	-	+	-	-	P2	-
pant44	363032	363166	362858	363277	135	PA0322	PA0323	<<<	174	111	-	+	-	+	8.8	-	-
pant45	367249	367433	367130	367456	185	PA0326	PA0327	<<>	119	23	-	+	+	-	-	-	-
pant46	371843	371948	371834	372090	106	PA0330	PA0331	<?>	9	142	-	+	+	-	-	-	-
pant47	397937	398067	397896	398223	131	PA0353	PA0354	<><	41	156	-	-	+	+	4.4	-	-
pant48	412084	412328	411872	412328	245	PA0367	PA0368	>><	212	0	-	+	+	-	-	-	-
pant49	413655	413808	413655	413932	154	PA0369	PA0370	<<<	0	124	-	-	+	-	-	-	-

pant50	449389	449483	449385	449638	95	PA0407	PA0408	<?>	4	155	-	+	-	-	-	-	-
pant51	484173	484273	484116	484403	101	PA0432	PA0433	<?>	57	130	-	-	+	-	-	AH riboswitch	-
pant52	501123	501317	501121	501375	195	PA0445	PA0446	<?<	2	58	-	+	-	+	2.0	-	-
pant53	507452	507589	507252	507630	138	PA0450	PA0451	<>>	200	41	-	+	+	+	3.2	-	-
pant54	514536	514620	514428	514774	85	PA0455	PA0456	<>>	108	154	-	-	+	-	-	-	-
pant55	517652	517740	517463	518082	89	PA0458	PA0459	<>>	189	342	-	-	+	+	246.3	-	-
pant56	521517	521870	521316	522464	354	PA0460	PA0462	>?>	201	594	+	+	-	+	3.1	-	-
pant57	521960	522172	521316	522464	213	PA0460	PA0462	>?>	644	292	+	+	+	-	-	-	++
pant58	522650	523205	522518	523253	556	PA0461	PA0463	<?>	132	48	+	-	-	+	5.3	-	+
pant59	522753	522868	522518	523253	116	PA0461	PA0463	<>>	235	385	-	-	+	+	8.4	-	++
pant60	522900	523037	522518	523253	138	PA0461	PA0463	<>>	382	216	-	+	+	+	9.9	-	+
pant61	545931	546048	545645	546333	118	PA0484	PA0485	>?<	286	285	-	+	-	+	0.4	-	-
pant62	546131	546304	545645	546333	174	PA0484	PA0485	>><	486	29	+	+	+	-	-	-	-
pant63	559211	559544	559079	559643	334	PA0499	PA0500	<?>	132	99	+	+	+	-	-	-	-
pant64	600177	600277	600177	600425	101	PA0542	PA0543	<<>	0	148	-	+	+	-	-	-	-
pant65	627957	628275	627766	628334	319	PA0572	PA0573	<><	191	59	-	+	+	-	-	-	-
pant66	648698	648847	648653	648930	150	PA0588	PA0589	<<<	45	83	-	+	+	+	7.0	-	-
pant67	674043	674369	673962	674418	327	PA0611	PA0612	<>>	81	49	-	+	+	-	-	-	-
pant68	698707	698894	698697	698931	188	PA0641	PA0642	>>>	10	37	-	+	+	-	-	-	-
pant69	703216	703328	703059	703476	113	PA0648	PA0649	>>>	157	148	-	+	-	-	-	-	-
pant70	706691	706815	706673	706943	125	PA0652	PA0653	<<>	18	128	-	-	+	-	-	-	-
pant71	707436	707677	707367	707685	242	PA0653	PA0654	>>>	69	8	-	+	+	+	0.3	-	-
pant72	727470	727602	727256	727607	133	PA0668.5	PA0669	>><	214	5	-	+	+	+	0.5	-	-
pant73	773742	773924	773697	774415	183	PA0701	PA0702	>><	45	491	-	+	+	-	-	-	-
pant74	781147	781246	780974	781258	100	PA0707	PA0708	<<<	173	12	-	+	-	+	2.3	-	-
pant75	783859	784028	783834	784172	170	PA0712	PA0713	<<>	25	144	-	-	+	-	-	-	-
pant76	785731	785832	785571	785968	102	PA0714.1	PA0715	>?>	160	136	-	+	+	-	-	-	-
pant77	788390	789051	788254	789143	662	PA0716	PA0717	>?>	136	92	-	+	+	-	-	-	+
pant78	791239	791330	791235	791369	92	PA0723	PA0724	>>>	4	39	-	+	-	-	-	-	-
pant79	796855	797228	796777	797250	374	PA0728	PA0729	>>>	78	22	-	-	+	-	-	-	-
pant80	798893	799149	798828	799278	257	PA0730	PA0731	<<<	65	129	-	+	+	-	-	-	-
pant81	804295	804574	804149	804655	280	PA0736	PA0737	<>>	146	81	-	+	+	-	-	-	-
pant82	804581	804652	804149	804655	72	PA0736	PA0737	<?>	432	3	-	+	+	-	-	-	-
pant83	828360	828477	828345	828616	118	PA0759	PA0760	<>>	15	139	-	+	+	+	2.8	-	-
pant84	845182	845255	845180	845277	74	PA0777	PA0778	<<<	2	22	-	-	+	-	-	-	-
pant85	851613	851745	851320	851782	133	PA0781	PA0782	<>>	293	37	-	-	+	+	0.0	-	-
pant86	855071	855234	854966	855276	164	PA0782	PA0783	>>>	105	42	-	+	+	-	-	-	-
pant87	863428	863611	863301	864094	184	PA0788	PA0789	<>>	127	483	-	+	+	-	-	-	-
pant88	863878	864029	863301	864094	152	PA0788	PA0789	<>>	577	65	-	+	+	-	-	-	-
pant89	883433	883594	883217	883776	162	PA0805	PA0806	<?>	216	182	-	-	+	-	-	-	-
pant90	884351	884629	884173	884798	279	PA0806	PA0807	>?>	178	169	+	-	-	-	-	-	-
pant91	892821	892966	892821	893040	146	PA0814	PA0815	<<>	0	74	-	+	+	-	-	-	-
pant92	896243	896374	896118	896415	132	PA0819	PA0820	>>>	125	41	-	-	+	+	6.1	-	-
pant93	913662	913761	913572	913776	100	PA0837	PA0838	>>>	90	15	-	+	-	-	-	-	-
pant94	916156	916300	916156	916345	145	PA0840	PA0841	>?>	0	45	-	+	+	-	-	-	-
pant95	927232	927551	927215	927879	320	PA0847	PA0849	<?>	17	328	-	-	+	+	6.9	-	++
pant96	927603	927784	927215	927879	182	PA0847	PA0849	<>>	388	95	-	+	+	-	-	-	++
pant97	936354	936454	936295	936458	101	PA0857	PA0858	>>>	59	4	-	+	+	-	-	-	-
pant98	946960	947064	946908	947204	105	PA0865	PA0866	>>>	52	140	-	+	-	+	5.2	-	-
pant99	950713	950898	950649	950961	186	PA0869	PA0870	<?<	64	63	-	+	+	-	-	-	-
pant100	999767	999920	999767	1000012	154	PA0916	PA0917	<?>	0	92	-	+	+	-	-	-	-
pant101	1006635	1006711	1006612	1006859	77	PA0920	PA0921	<<>	23	148	-	+	-	+	7.5	-	-
pant102	1038894	1038971	1038886	1039239	78	PA0951	PA0952	>><	8	268	-	+	-	-	-	-	-
pant103	1045522	1045656	1045315	1045831	135	PA0958	PA0959	<<>	207	175	-	+	+	+	2.8	-	-

pant104	1047055	1047183	1046912	1047548	129	PA0961	PA0962	<<<	143	365	+	+	+	-	-	-	-
pant105	1047198	1047439	1046912	1047548	242	PA0961	PA0962	<<<	286	109	+	+	+	-	-	-	-
pant106	1063787	1064073	1063545	1064554	287	PA0981	PA0982	>?<	242	481	-	+	-	+	3.0	-	-
pant107	1064226	1064534	1063545	1064554	309	PA0981	PA0982	>?<	681	20	+	+	+	-	-	-	-
pant108	1070187	1070289	1070174	1070293	103	PA0988	PA0989	>><	13	4	-	+	+	-	-	-	-
pant109	1078124	1078298	1077905	1078461	175	PA0995	PA0996	<>>	219	163	+	+	+	+	51.5	-	-
pant110	1087101	1087236	1087096	1087842	136	PA1003	PA1004	<>>	5	606	+	+	+	-	-	-	-
pant111	1087260	1087621	1087096	1087842	362	PA1003	PA1004	<<>	164	221	+	+	+	-	-	-	-
pant112	1099180	1099330	1099129	1099668	151	PA1015	PA1016	<>>	51	338	-	+	+	-	-	-	-
pant113	1105672	1105853	1105365	1105887	182	PA1019	PA1020	>?>	307	34	-	+	+	+	3.8	-	-
pant114	1116110	1116211	1116061	1116212	102	PA1028	PA1029	>>>	49	1	-	+	+	-	-	-	-
pant115	1117635	1117968	1117610	1118157	334	PA1030.1	PA1031	>>>	25	189	-	+	+	+	0.3	-	-
pant116	1135225	1135406	1135203	1135407	182	PA1047	PA1048	>>>	22	1	-	-	+	-	-	-	-
pant117	1141746	1142015	1141733	1142060	270	PA1053	PA1054	>?>	13	45	-	+	+	-	-	-	-
pant118	1163013	1163160	1162978	1163275	148	PA1074	PA1075	<>>	35	115	-	-	+	+	3.2	-	-
pant119	1182779	1182921	1182698	1183057	143	PA1091	PA1092	><>	81	136	-	-	+	+	3.5	-	-
pant120	1204789	1205025	1204782	1205030	237	PA1112	PA1112.1	>><	7	5	+	+	+	+	3.0	-	++
pant121	1205337	1205617	1205331	1205770	281	PA1112.1	PA1113	<>>	6	153	-	+	+	-	-	-	-
pant122	1246048	1246223	1245929	1246443	176	PA1151	PA1152	>>>	119	220	-	-	+	-	-	-	-
pant123	1248655	1248804	1248487	1249018	150	PA1153	PA1154	<?<	168	214	-	+	+	+	2.8	-	-
pant124	1248876	1248985	1248487	1249018	110	PA1153	PA1154	<><	389	33	-	+	-	-	-	-	-
pant125	1254329	1254720	1254310	1255041	392	PA1156	PA1157	<<>	19	321	+	+	+	+	0.2	-	-
pant126	1254790	1254903	1254310	1255041	114	PA1156	PA1157	<<>	480	138	+	+	+	+	0.2	-	-
pant127	1255801	1255974	1255753	1256093	174	PA1157	PA1158	>>>	48	119	-	+	+	-	-	-	+
pant128	1266929	1267100	1266783	1267281	172	PA1167	PA1168	><>	146	181	-	+	-	+	5.3	-	-
pant129	1270848	1270960	1270516	1270971	113	PA1170	PA1171	<>>	332	11	-	+	+	-	-	-	-
pant130	1284338	1284437	1284178	1284512	100	PA1182	PA1183	<>>	160	75	-	+	-	-	-	-	-
pant131	1292441	1292545	1292413	1292570	105	PA1190	PA1191	<?<	28	25	-	+	+	+	6.4	-	+
pant132	1324155	1324426	1324110	1324792	272	PA1221	PA1222	<<<	45	366	-	+	+	+	45.4	-	-
pant133	1348429	1348618	1348402	1349415	190	PA1244	PA1245	<<>	27	797	-	-	+	-	-	C4	-
pant134	1348836	1349074	1348402	1349415	239	PA1244	PA1245	<<>	434	341	-	+	+	-	-	-	-
pant135	1349247	1349393	1348402	1349415	147	PA1244	PA1245	<?>	845	22	-	+	+	+	8.4	-	-
pant136	1355391	1355625	1355273	1355630	235	PA1248	PA1249	>>>	118	5	-	+	+	+	3.4	-	-
pant137	1357149	1357257	1357071	1357316	109	PA1249	PA1250	>>>	78	59	-	-	+	+	10.1	-	-
pant138	1372693	1372804	1372683	1372978	112	PA1263	PA1264	>?<	10	174	-	-	+	-	-	-	++
pant139	1379227	1379357	1379169	1379412	131	PA1269	PA1270	<>>	58	55	-	-	+	+	4.3	-	-
pant140	1381509	1381765	1381453	1381803	257	PA1270	PA1271	>>>	56	38	-	-	+	-	-	Cobalamin	+
pant141	1400537	1400648	1400506	1400710	112	PA1288	PA1289	<><	31	62	-	-	+	-	-	-	-
pant142	1406850	1406989	1406753	1407006	140	PA1295	PA1296	>>>	97	17	-	+	+	-	-	-	-
pant143	1427465	1428040	1427454	1428079	576	PA1316	PA1317	><>	11	39	+	+	+	-	-	-	-
pant144	1471942	1472164	1471673	1472546	223	PA1359	PA1361	<><	269	382	-	+	+	-	-	-	++
pant145	1474178	1474389	1473981	1474390	212	PA1361	PA1362	<<>	197	1	-	+	+	-	-	-	-
pant146	1479852	1480066	1479792	1480214	215	PA1366	PA1367	<<<	60	148	-	+	+	-	-	-	-
pant147	1482904	1482991	1482759	1483122	88	PA1368	PA1369	<<>	145	131	-	+	-	+	10.9	-	-
pant148	1485764	1485865	1485764	1486265	102	PA1370	PA1371	>><	0	400	-	-	+	+	5.6	-	-
pant149	1486012	1486265	1485764	1486265	254	PA1370	PA1371	>?<	248	0	-	-	+	-	-	-	-
pant150	1489164	1489362	1489096	1489814	199	PA1372	PA1373	<<<	68	452	-	-	+	-	-	-	-
pant151	1489557	1489652	1489096	1489814	96	PA1372	PA1373	<><	461	162	-	+	+	-	-	-	-
pant152	1602058	1602160	1601894	1602178	103	PA1474	PA1475	<>>	164	18	-	+	+	-	-	-	-
pant153	1609295	1609403	1609292	1609737	109	PA1483	PA1484	>><	3	334	-	+	+	-	-	-	-
pant154	1617190	1617423	1617086	1617488	234	PA1489	PA1490	<<<	104	65	-	+	-	+	2.7	-	-
pant155	1618312	1618442	1618266	1618611	131	PA1490	PA1491	<<>	46	169	-	+	+	-	-	-	-
pant156	1625863	1625937	1625696	1626061	75	PA1497	PA1498	<<<	167	124	-	+	-	-	-	-	-
pant157	1644795	1645115	1644726	1645210	321	PA1512	PA1513	<><	69	95	-	+	+	-	-	-	-

pant158	1681845	1681933	1681806	1682050	89	PA1544	PA1545	<?>	39	117	-	-	+	+	8.2	-	-
pant159	1703072	1703249	1703053	1703378	178	PA1562	PA1563	<?>	19	129	-	+	+	+	4.0	-	-
pant160	1717940	1718216	1717895	1718385	277	PA1578	PA1579	>>>	45	169	+	+	+	+	0.2	-	++
pant161	1729853	1730100	1729853	1730180	248	PA1587	PA1588	>>>	0	80	-	-	+	-	-	-	-
pant162	1732291	1732390	1732235	1732544	100	PA1589	PA1590	>?>	56	154	-	+	-	-	-	-	-
pant163	1758602	1758698	1758598	1758747	97	PA1613	PA1614	<?>	4	49	-	+	+	-	-	-	-
pant164	1773494	1773592	1773441	1773681	99	PA1629	PA1630	<?>	53	89	-	+	-	-	-	-	-
pant165	1802487	1802745	1802454	1803625	259	PA1655	PA1656	>?>	33	880	-	+	+	-	-	-	-
pant166	1802750	1803294	1802454	1803625	545	PA1655	PA1656	>>>	296	331	-	+	+	+	8.8	-	-
pant167	1803380	1803604	1802454	1803625	225	PA1655	PA1656	>>>	926	21	-	+	-	-	-	-	-
pant168	1824753	1824900	1824724	1824968	148	PA1672	PA1673	<<>	29	68	-	-	+	-	-	-	-
pant169	1882442	1882555	1882344	1882594	114	PA1737	PA1738	>><	98	39	-	-	+	-	-	-	-
pant170	1886040	1886112	1886034	1886277	73	PA1741	PA1742	<<>	6	165	-	-	+	+	4.0	-	-
pant171	1904315	1904419	1904248	1904478	105	PA1760	PA1761	>>>	67	59	-	-	+	+	8.0	-	-
pant172	1947992	1948157	1947805	1948501	166	PA1796.4	PA1797	>><	187	344	-	+	+	-	-	-	-
pant173	1948239	1948395	1947805	1948501	157	PA1796.4	PA1797	>?<	434	106	-	+	-	+	2.2	-	-
pant174	1988953	1989087	1988852	1989108	135	PA1829	PA1830	<><	101	21	-	+	+	-	-	-	-
pant175	2005007	2005107	2004893	2005253	101	PA1845	PA1846	<<>	114	146	-	+	+	+	2.8	-	-
pant176	2012651	2012770	2012531	2012814	120	PA1852	PA1853	<?>	120	44	-	+	+	-	-	-	-
pant177	2016813	2016924	2016582	2016927	112	PA1856	PA1857	<><	231	3	-	-	+	-	-	-	-
pant178	2028243	2028416	2028014	2028453	174	PA1866	PA1867	>>>	229	37	-	+	+	-	-	-	-
pant179	2031706	2031839	2031706	2031987	134	PA1869	PA1870	>>>	0	148	-	+	-	+	4.2	-	-
pant180	2068748	2068961	2068729	2069489	214	PA1897	PA1898	<?>	19	528	-	+	-	+	12.0	-	-
pant181	2068964	2069268	2068729	2069489	305	PA1897	PA1898	<>>	235	221	-	+	+	-	-	-	-
pant182	2070481	2070599	2070204	2070684	119	PA1898	PA1899	>>>	277	85	-	+	+	+	62.3	-	-
pant183	2093320	2093418	2093251	2093443	99	PA1918	PA1919	<><	69	25	-	+	-	-	-	-	-
pant184	2104152	2104333	2104097	2104596	182	PA1925	PA1926	>>>	55	263	-	+	+	-	-	-	++
pant185	2104397	2104580	2104097	2104596	184	PA1925	PA1926	>>>	300	16	-	+	+	+	0.3	-	++
pant186	2118138	2118426	2118098	2118584	289	PA1936	PA1937	><>	40	158	-	+	+	+	3.0	-	-
pant187	2119868	2120185	2119748	2120224	318	PA1938	PA1939	>><	120	39	-	+	+	+	2.7	-	-
pant188	2130677	2130840	2130636	2130853	164	PA1945	PA1946	><>	41	13	-	+	+	-	-	-	-
pant189	2153211	2153381	2153069	2153593	171	PA1971	PA1972	<>>	142	212	-	+	+	-	-	-	-
pant190	2189965	2190071	2189884	2190083	107	PA2002	PA2003	>><	81	12	-	+	+	-	-	-	-
pant191	2230425	2230600	2230184	2230688	176	PA2039	PA2040	<>>	241	88	-	+	-	+	3.2	-	-
pant192	2232100	2232186	2232066	2232221	87	PA2040	PA2041	>>>	34	35	-	+	+	+	4.7	-	-
pant193	2233636	2233805	2233593	2234079	170	PA2041	PA2042	>>>	43	274	-	+	+	-	-	-	-
pant194	2282233	2282379	2281579	2282479	147	PA2075	PA2076	<?>	654	100	-	+	+	-	-	-	-
pant195	2291596	2291763	2291590	2291790	168	PA2082	PA2083	>?>	6	27	-	+	-	-	-	-	-
pant196	2314392	2314498	2314250	2314511	107	PA2102	PA2103	>>>	142	13	-	+	-	+	4.6	-	-
pant197	2330441	2330688	2330425	2330958	248	PA2118	PA2119	>?<	16	270	+	+	+	-	-	-	-
pant198	2330834	2330934	2330425	2330958	101	PA2118	PA2119	>?<	409	24	-	-	+	+	4.7	-	-
pant199	2341653	2341988	2341641	2342492	336	PA2127	PA2128	<<>	12	504	-	+	+	+	3.7	-	-
pant200	2341875	2341978	2341641	2342492	104	PA2127	PA2128	<>>	234	514	-	+	-	+	3.4	-	-
pant201	2342040	2342162	2341641	2342492	123	PA2127	PA2128	<<>	399	330	-	+	-	-	-	-	-
pant202	2396383	2396488	2396253	2396535	106	PA2174	PA2175	<><	130	47	-	-	+	+	4.2	-	-
pant203	2442721	2443060	2442692	2443160	340	PA2220	PA2221	>?>	29	100	+	+	+	-	-	-	-
pant204	2444367	2444624	2444367	2444885	258	PA2221	PA2222	>><	0	261	-	+	+	-	-	-	-
pant205	2447361	2447521	2447326	2447572	161	PA2224	PA2225	<<<	35	51	-	+	+	+	3.5	-	-
pant206	2450801	2450900	2450766	2451706	100	PA2228	PA2229	<<>	35	806	-	+	-	-	-	-	-
pant207	2451228	2451464	2450766	2451706	237	PA2228	PA2229	<?>	462	242	+	-	-	+	2.2	-	-
pant208	2453289	2453505	2453141	2453666	217	PA2230	PA2231	><>	148	161	-	+	+	-	-	-	-
pant209	2493101	2493210	2493051	2493217	110	PA2263	PA2264	>?>	50	7	-	+	+	+	9.3	-	-
pant210	2521175	2521247	2521105	2521257	73	PA2290	PA2291	<?<	70	10	-	-	+	-	-	-	-
pant211	2522623	2522747	2522617	2522957	125	PA2291	PA2292	<<<	6	210	-	+	+	-	-	-	-

pant212	2531867	2531979	2531841	2532058	113	PA2300	PA2301	<<<	26	79	-	+	+	-	-	-	-
pant213	2557983	2558363	2557965	2558917	381	PA2319	PA2320	>><	18	554	+	+	+	-	-	-	-
pant214	2559995	2560132	2559950	2560143	138	PA2320	PA2321	<>>	45	11	-	+	+	-	-	-	-
pant215	2614740	2614850	2614685	2614892	111	PA2364	PA2365	<>>	55	42	-	+	+	+	3.3	-	-
pant216	2623966	2624200	2623857	2624203	235	PA2372	PA2373	>?>	109	3	-	+	-	+	4.1	-	-
pant217	2627303	2627451	2627176	2627451	149	PA2375	PA2376	<>>	127	0	+	+	+	-	-	-	-
pant218	2629644	2629914	2629500	2629915	271	PA2377	PA2378	>><	144	1	-	+	+	-	-	-	-
pant219	2687349	2687474	2687179	2687496	126	PA2402	PA2403	<?>	170	22	-	+	-	+	5.9	-	-
pant220	2705330	2705770	2705298	2705772	441	PA2421	PA2422	>><>	32	2	+	+	+	-	-	-	-
pant221	2707261	2707595	2707182	2707665	335	PA2423	PA2424	>><	79	70	+	+	+	-	-	-	-
pant222	2728318	2728427	2728142	2728541	110	PA2431	PA2432	<<>	176	114	-	+	+	-	-	-	-
pant223	2737455	2737705	2737195	2737881	251	PA2439	PA2440	<<>	260	176	-	+	+	+	3.4	-	-
pant224	2760623	2760747	2760620	2760871	125	PA2460	PA2461	<?<	3	124	-	+	+	-	-	-	-
pant225	2761392	2761657	2761352	2761920	266	PA2461	PA2462	<<<	40	263	-	-	+	+	0.3	-	-
pant226	2761671	2761919	2761352	2761920	249	PA2461	PA2462	<<<	319	1	-	-	+	-	-	-	-
pant227	2822388	2822488	2822323	2822573	101	PA2504	PA2505	<<<	65	85	-	-	+	+	24.6	-	-
pant228	2843524	2843740	2843306	2843817	217	PA2522	PA2523	<>>	218	77	-	+	+	-	-	-	-
pant229	2893863	2894130	2893829	2894451	268	PA2559	PA2560	<?<	34	321	-	+	+	-	-	-	-
pant230	2916938	2917154	2916749	2917171	217	PA2580	PA2581	<<>	189	17	-	+	+	+	2.4	-	-
pant231	2943992	2944212	2943916	2944254	221	PA2600	PA2601	<?<	76	42	-	+	+	-	-	-	-
pant232	2959289	2959461	2959241	2959467	173	PA2615	PA2616	<>>	48	6	-	-	+	-	-	-	-
pant233	2964966	2965175	2964844	2965200	210	PA2621	PA2622	<<>	122	25	-	-	+	-	-	-	-
pant234	2967068	2967160	2966803	2967160	93	PA2623	PA2624	<>>	265	0	-	+	-	+	0.4	-	-
pant235	2977424	2977601	2977162	2977755	178	PA2633	PA2634	>>>	262	154	-	+	+	+	3.1	-	-
pant236	2982246	2982481	2982182	2982780	236	PA2636	PA2637	>><>	64	299	-	+	+	+	2.5	-	-
pant237	2982583	2982654	2982182	2982780	72	PA2636	PA2637	>><>	401	126	-	+	-	+	2.9	-	-
pant238	3008717	3008809	3008633	3008846	93	PA2659	PA2660	<<>	84	37	-	+	+	-	-	-	-
pant239	3017012	3017103	3016885	3017233	92	PA2668	PA2669	<<<	127	130	-	+	-	-	-	-	-
pant240	3035720	3036327	3035703	3037885	608	PA2684	PA2686	<?>	17	1558	-	+	+	-	-	-	+
pant241	3036726	3037186	3035703	3037885	461	PA2684	PA2686	<<>	1023	699	-	+	+	-	-	-	+
pant242	3037237	3037426	3035703	3037885	190	PA2684	PA2686	<<>	1534	459	-	+	+	-	-	-	++
pant243	3037611	3037811	3035703	3037885	201	PA2684	PA2686	<?>	1908	74	-	+	+	-	-	-	-
pant244	3044769	3044839	3044767	3044946	71	PA2690	PA2691	<?<	2	107	-	+	-	-	-	-	-
pant245	3050173	3050313	3050162	3050330	141	PA2696	PA2697	<><	11	17	-	+	+	-	-	-	-
pant246	3065879	3066090	3065877	3066295	212	PA2711	PA2712	><>	2	205	-	+	+	-	-	-	-
pant247	3067747	3067961	3067687	3068039	215	PA2713	PA2714	<<>	60	78	-	+	+	-	-	-	-
pant248	3087529	3088155	3087492	3088659	627	PA2729	PA2730	>><	37	504	-	+	+	-	-	-	-
pant249	3088213	3088434	3087492	3088659	222	PA2729	PA2730	>?<	721	225	-	+	+	-	-	-	-
pant250	3088452	3088618	3087492	3088659	167	PA2729	PA2730	>><	960	41	-	+	-	-	-	-	-
pant251	3094233	3094534	3094186	3094756	302	PA2733	PA2734	<?<	47	222	-	+	+	-	-	-	++
pant252	3098983	3099074	3098980	3099305	92	PA2736	PA2736.1	<?<	3	231	-	+	+	-	-	-	-
pant253	3108450	3108697	3108384	3108969	248	PA2746	PA2747	>><	66	272	+	+	+	-	-	-	-
pant254	3108698	3108907	3108384	3108969	210	PA2746	PA2747	>><	314	62	+	+	+	-	-	-	-
pant255	3109522	3109714	3109258	3109742	193	PA2747	PA2748	<<<	264	28	-	+	+	-	-	-	-
pant256	3110558	3110777	3110526	3110900	220	PA2748	PA2749	<?>	32	123	+	+	+	-	-	-	-
pant257	3115808	3115984	3115634	3116653	177	PA2754	PA2755	><>	174	669	-	+	-	+	3.4	-	-
pant258	3116209	3116365	3115634	3116653	157	PA2754	PA2755	>?>	575	288	-	+	+	+	2.6	-	-
pant259	3119820	3119930	3119709	3120072	111	PA2759	PA2760	<<>	111	142	-	+	+	+	2.7	-	-
pant260	3122766	3122846	3122587	3123598	81	PA2763	PA2764	<<>	179	752	-	-	+	+	5.8	-	-
pant261	3123173	3123434	3122587	3123598	262	PA2763	PA2764	<<>	586	164	+	+	+	-	-	-	-
pant262	3129277	3129477	3129072	3129728	201	PA2770	PA2771	<<>	205	251	-	+	+	-	-	-	-
pant263	3129461	3129635	3129072	3129728	175	PA2770	PA2771	<<>	389	93	-	+	+	-	-	-	-
pant264	3138591	3138791	3138533	3139010	201	PA2781	PA2782	>>>	58	219	-	+	+	+	6.0	-	-
pant265	3150508	3150730	3150435	3150885	223	PA2793	PA2794	<<<	73	155	+	+	+	-	-	-	-

pant266	3193555	3193835	3193520	3193885	281	PA2839	PA2840	>?<	35	50	+	+	+	-	-	-	-
pant267	3195665	3195766	3195590	3195899	102	PA2840	PA2841	<<>	75	133	-	-	+	+	0.2	-	-
pant268	3218534	3218648	3218420	3218665	115	PA2865	PA2866	<?>	114	17	-	-	+	-	-	-	-
pant269	3249634	3249821	3249632	3249976	188	PA2894	PA2895	<<<	2	155	-	+	+	-	-	-	-
pant270	3251354	3251454	3251323	3251650	101	PA2896	PA2897	<<<	31	196	-	-	+	+	2.3	-	-
pant271	3251355	3251627	3251323	3251650	273	PA2896	PA2897	<<<	32	23	-	+	-	-	-	-	-
pant272	3261207	3261426	3261186	3261521	220	PA2906	PA2907	<<>	21	95	-	+	+	-	-	Cobalamin	-
pant273	3265500	3265795	3265212	3265847	296	PA2910	PA2911	<>>	288	52	+	+	+	-	-	Cobalamin	+
pant274	3281758	3281897	3281667	3281987	140	PA2926	PA2927	><>	91	90	-	+	-	-	-	-	-
pant275	3284713	3284940	3284599	3285006	228	PA2928	PA2929	<<<	114	66	-	+	+	-	-	-	-
pant276	3293795	3294072	3293782	3294281	278	PA2938	PA2939	<?>	13	209	+	+	+	-	-	-	-
pant277	3305612	3305787	3305565	3305918	176	PA2945	PA2946	<<>	47	131	-	+	+	-	-	Cobalamin	++
pant278	3316654	3316731	3316651	3316799	78	PA2956	PA2957	<<<	3	68	-	-	+	-	-	-	-
pant279	3346360	3346486	3345797	3346976	127	PA2987	PA2989	<<>	563	490	-	+	+	-	-	-	+
pant280	3347047	3347276	3347040	3348072	230	PA2988	PA2990	<<>	7	796	-	+	+	-	-	-	-
pant281	3366842	3366917	3366755	3366969	76	PA3005	PA3006	<<<	87	52	-	-	+	+	0.3	-	-
pant282	3376636	3376826	3376608	3377034	191	PA3014	PA3015	<>>	28	208	-	+	+	-	-	-	+
pant283	3378402	3378478	3378305	3378511	77	PA3016	PA3017	<<>	97	33	-	+	-	+	3.8	-	-
pant284	3409983	3410069	3409954	3410263	87	PA3046	PA3047	<<>	29	194	-	+	+	-	-	-	-
pant285	3414160	3414264	3413941	3414400	105	PA3048	PA3049	<?>	219	136	-	+	-	+	0.4	-	-
pant286	3415758	3416030	3415730	3416060	273	PA3050	PA3052	<>>	28	30	-	+	+	-	-	-	+
pant287	3416582	3416725	3416068	3417112	144	PA3051	PA3053	<>>	514	387	-	+	+	-	-	-	+
pant288	3441450	3441689	3441441	3441712	240	PA3068	PA3069	<<<	9	23	-	+	+	+	4.8	-	-
pant289	3442446	3442661	3442331	3442822	216	PA3069	PA3070	<<>	115	161	-	+	+	+	3.5	-	-
pant290	3459955	3460268	3459934	3460379	314	PA3082	PA3083	<<<	21	111	-	+	+	-	-	-	-
pant291	3475180	3475456	3475171	3475456	277	PA3094	PA3094.1	>?<	9	0	-	+	+	-	-	-	-
pant292	3545653	3545831	3545075	3545880	179	PA3159	PA3160	<?<	578	49	-	+	-	+	3.1	-	-
pant293	3567471	3567603	3567378	3567605	133	PA3176	PA3177	>?>	93	2	-	+	-	-	-	-	-
pant294	3581773	3581852	3581647	3582170	80	PA3190	PA3191	<><	126	318	-	+	+	-	-	-	-
pant295	3581900	3582070	3581647	3582170	171	PA3190	PA3191	<<<	253	100	-	+	+	-	-	-	-
pant296	3588569	3588705	3588438	3588740	137	PA3195	PA3196	>><	131	35	-	+	-	+	2.9	-	-
pant297	3592929	3593104	3592929	3593143	176	PA3200	PA3201	<<>	0	39	-	-	+	-	-	-	-
pant298	3612366	3612474	3612360	3612558	109	PA3224	PA3225	<>>	6	84	-	-	+	+	4.1	-	-
pant299	3616919	3617018	3616763	3617043	100	PA3228	PA3229	>>>	156	25	-	+	+	+	2.8	-	-
pant300	3630634	3630737	3630606	3630764	104	PA3242	PA3243	<<>	28	27	-	+	+	+	5.0	-	-
pant301	3648952	3649078	3648917	3649704	127	PA3261	PA3262	>><	35	626	-	+	+	-	-	-	+
pant302	3649100	3649605	3648917	3649704	506	PA3261	PA3262	>><	183	99	+	+	+	-	-	-	-
pant303	3669652	3669788	3669589	3669820	137	PA3276	PA3277	<?<	63	32	-	+	+	-	-	-	-
pant304	3677575	3677718	3677082	3677719	144	PA3284	PA3285	<<<	493	1	-	+	+	-	-	P17	+
pant305	3679417	3679810	3679387	3679945	394	PA3286	PA3287	<<>	30	135	-	+	+	-	-	-	-
pant306	3706855	3707061	3706855	3707089	207	PA3307	PA3308	<?>	0	28	-	+	+	-	-	-	-
pant307	3722785	3722865	3722760	3722989	81	PA3319	PA3320	<<<	25	124	-	+	+	+	3.5	-	-
pant308	3729139	3729327	3729128	3729470	189	PA3325	PA3326	>><	11	143	-	-	+	+	6.2	-	-
pant309	3730189	3730383	3730077	3730556	195	PA3326	PA3327	<<>	112	173	-	+	-	+	17.7	-	-
pant310	3730416	3730556	3730077	3730556	141	PA3326	PA3327	<>>	339	0	-	+	+	+	47.2	-	-
pant311	3751917	3752011	3751768	3752043	95	PA3340	PA3341	<?<	149	32	-	-	+	+	4.8	-	-
pant312	3759712	3759942	3759682	3759995	231	PA3347	PA3348	<<<	30	53	-	+	+	-	-	-	-
pant313	3772588	3772694	3772562	3773028	107	PA3360	PA3361	<<>	26	334	-	+	-	-	-	-	-
pant314	3772905	3773012	3772562	3773028	108	PA3360	PA3361	<>>	343	16	-	-	+	+	81.9	-	-
pant315	3803431	3803595	3803344	3803616	165	PA3397	PA3398	<>>	87	21	-	+	+	-	-	-	-
pant316	3819826	3819939	3819797	3820005	114	PA3412	PA3413	<>>	29	66	-	+	+	+	3.4	-	-
pant317	3841795	3842086	3841738	3842273	292	PA3433	PA3434	>><	57	187	-	+	+	-	-	-	-
pant318	3843293	3843364	3843291	3843652	72	PA3434	PA3435	<?<	2	288	-	+	-	-	-	-	-
pant319	3843449	3843551	3843291	3843652	103	PA3434	PA3435	<?<	158	101	-	+	-	-	-	-	-

pant320	3865152	3865261	3865127	3865280	110	PA3457	PA3458	>?<	25	19	-	-	+	+	2.4	-	-
pant321	3877915	3878018	3877912	3878032	104	PA3465	PA3466	>><	3	14	-	-	+	-	-	-	-
pant322	3885259	3885449	3885255	3885710	191	PA3471	PA3472	<<>	4	261	-	+	+	-	-	-	-
pant323	3885464	3885686	3885255	3885710	223	PA3471	PA3472	<>>	209	24	-	-	+	-	-	-	-
pant324	3893030	3893253	3893009	3893431	224	PA3479	PA3480	<<<	21	178	-	+	+	+	4.2	-	+
pant325	3900073	3900231	3899939	3900232	159	PA3485	PA3486	>>>	134	1	-	+	+	-	-	-	-
pant326	3930268	3930716	3930231	3930742	449	PA3514	PA3515	<><	37	26	-	+	+	-	-	-	-
pant327	3936883	3937150	3936838	3937237	268	PA3519	PA3520	<<<	45	87	+	+	-	+	6.5	-	-
pant328	3937547	3937816	3937433	3938019	270	PA3520	PA3521	<><	114	203	-	+	+	-	-	-	-
pant329	3949697	3949817	3949594	3949852	121	PA3529	PA3530	<<>	103	35	-	+	+	-	-	-	-
pant330	3954610	3954687	3954609	3954906	78	PA3534	PA3535	<<>	1	219	-	-	+	+	5.9	-	-
pant331	3957895	3957978	3957895	3958287	84	PA3535	PA3536	>><	0	309	-	+	-	-	-	-	-
pant332	3958050	3958149	3957895	3958287	100	PA3535	PA3536	>><	102	138	-	+	+	-	-	-	-
pant333	3958814	3959028	3958756	3959032	215	PA3536	PA3537	<<>	58	4	+	+	+	-	-	-	-
pant334	3979483	3979687	3979477	3979859	205	PA3551	PA3552	><>	6	172	-	+	+	-	-	-	-
pant335	4009052	4009156	4009052	4009541	105	PA3577	PA3578	<<<	0	385	-	-	+	-	-	-	+
pant336	4009456	4009526	4009052	4009541	71	PA3577	PA3578	<?<	404	15	-	+	-	-	-	-	-
pant337	4012689	4012804	4012542	4012805	116	PA3580	PA3581	<?>	147	1	-	+	-	+	0.1	-	-
pant338	4029599	4029753	4029541	4029758	155	PA3594	PA3595	>>>	58	5	-	+	+	-	-	-	-
pant339	4064994	4065078	4064988	4065118	85	PA3629	PA3630	<<>	6	40	-	+	-	+	3.3	-	-
pant340	4072509	4072650	4072489	4072658	142	PA3637	PA3638	<<<	20	8	-	+	+	-	-	-	-
pant341	4080262	4080358	4080224	4080439	97	PA3641	PA3642	<?<	38	81	-	+	-	+	0.2	-	-
pant342	4092998	4093157	4092969	4093166	160	PA3654	PA3655	<<<	29	9	-	+	+	+	0.2	-	-
pant343	4094917	4095015	4094908	4095171	99	PA3656	PA3657	<<>	9	156	-	-	+	-	-	t44	+
pant344	4133987	4134089	4133759	4134137	103	PA3690	PA3691	>>>	228	48	-	+	+	-	-	-	-
pant345	4160036	4160301	4160006	4160328	266	PA3713	PA3714	><>	30	27	-	+	+	+	2.1	-	-
pant346	4170519	4170697	4170484	4170769	179	PA3724	PA3725	<<<	35	72	-	+	+	+	16.3	-	-
pant347	4183467	4183589	4183229	4183709	123	PA3732	PA3733	<<>	238	120	-	+	+	-	-	-	-
pant348	4185172	4185429	4184940	4185521	258	PA3733	PA3734	><>	232	92	+	-	-	+	5.5	-	-
pant349	4189585	4189708	4189548	4189760	124	PA3736	PA3737	<<<	37	52	-	+	+	-	-	-	-
pant350	4241090	4241338	4240580	4241341	249	PA3781	PA3782	>><	510	3	-	+	+	-	-	-	-
pant351	4281929	4282070	4281752	4282077	142	PA3824.1	PA3825	>><	177	7	-	+	+	+	0.3	-	-
pant352	4288718	4288937	4288607	4288942	220	PA3830	PA3831	<?>	111	5	-	-	+	-	-	-	-
pant353	4295455	4296062	4294606	4297249	608	PA3835	PA3836	<<>	849	1187	+	-	-	-	-	-	-
pant354	4296539	4296803	4294606	4297249	265	PA3835	PA3836	<<>	1933	446	-	+	-	+	5.3	-	-
pant355	4296877	4296974	4294606	4297249	98	PA3835	PA3836	<<>	2271	275	-	-	+	+	9.1	-	-
pant356	4321074	4321255	4320821	4321433	182	PA3860	PA3861	<>>	253	178	-	+	+	-	-	-	-
pant357	4326945	4327123	4326396	4327696	179	PA3865	PA3866	<><	549	573	-	-	+	+	7.8	-	-
pant358	4344851	4344951	4344836	4345009	101	PA3877	PA3878	<>>	15	58	-	+	+	-	-	-	-
pant359	4352499	4352599	4352495	4352827	101	PA3886	PA3887	>><	4	228	-	+	+	+	5.9	-	+
pant360	4354229	4354375	4354103	4354543	147	PA3887	PA3888	<<>	126	168	-	+	+	-	-	-	-
pant361	4388737	4388953	4388728	4389230	217	PA3919	PA3920	<<>	9	277	-	+	+	-	-	-	+
pant362	4403499	4403704	4403427	4403706	206	PA3927	PA3928	>><	72	2	-	+	+	-	-	-	-
pant363	4406484	4406652	4406369	4406805	169	PA3930	PA3931	<>>	115	153	-	+	+	+	7.6	-	-
pant364	4418063	4418200	4417935	4418302	138	PA3939	PA3940	<>>	128	102	-	+	+	-	-	-	-
pant365	4433194	4433358	4433171	4433410	165	PA3952	PA3953	<<>	23	52	-	+	+	-	-	-	-
pant366	4472724	4472848	4472199	4473622	125	PA3991	PA3993	<<>	525	774	-	-	+	+	5.2	-	++
pant367	4472877	4473081	4472199	4473622	205	PA3991	PA3993	<>>	678	541	-	+	+	-	-	-	++
pant368	4474657	4474804	4474640	4474958	148	PA3993	PA3994	>><	17	154	-	+	+	-	-	-	+
pant369	4474837	4474958	4474640	4474958	122	PA3993	PA3994	>><	197	0	-	+	+	-	-	-	-
pant370	4514358	4514685	4514349	4514694	328	PA4033	PA4034	>>>	9	9	+	+	+	-	-	-	-
pant371	4518941	4519197	4518871	4519274	257	PA4036	PA4037	><>	70	77	+	+	+	-	-	-	-
pant372	4523842	4523939	4523754	4523984	98	PA4040	PA4041	>><	88	45	-	+	+	-	-	-	-
pant373	4536503	4536632	4536494	4536920	130	PA4055	PA4056	<<<	9	288	-	+	+	-	-	-	-

pant374	4545655	4545745	4545306	4545745	91	PA4067	PA4068	>><	349	0	-	+	+	-	-	-	-
pant375	4547578	4547720	4547553	4547820	143	PA4069	PA4070	<<<	25	100	-	+	-	-	-	-	-
pant376	4559680	4559844	4559651	4560294	165	PA4080	PA4081	<<<	29	450	-	+	+	-	-	-	-
pant377	4559856	4560204	4559651	4560294	349	PA4080	PA4081	<><	205	90	+	+	+	-	-	-	-
pant378	4592559	4592662	4592432	4592989	104	PA4108	PA4109	>><	127	327	-	+	+	+	4.1	-	-
pant379	4592808	4592973	4592432	4592989	166	PA4108	PA4109	>><	376	16	-	+	+	+	3.0	-	-
pant380	4601449	4601564	4601421	4601573	116	PA4113	PA4114	<<<	28	9	-	+	+	-	-	-	-
pant381	4629195	4629347	4629186	4629943	153	PA4138	PA4139	<<>	9	596	-	+	+	-	-	-	-
pant382	4632172	4632301	4632169	4632476	130	PA4140	PA4141	>><	3	175	-	+	+	+	2.7	-	-
pant383	4639348	4639448	4639292	4639501	101	PA4146	PA4147	>><	56	53	-	-	+	+	5.0	-	-
pant384	4648578	4648773	4648564	4648998	196	PA4154	PA4155	<><	14	225	-	+	+	-	-	-	-
pant385	4670823	4671197	4670785	4671318	375	PA4174	PA4175	>?>	38	121	-	+	+	+	3.9	-	-
pant386	4696088	4696278	4696036	4696280	191	PA4197	PA4198	<>>	52	2	-	+	-	-	-	-	-
pant387	4713137	4713370	4713100	4713795	234	PA4209	PA4210	<?>	37	425	-	-	+	+	200.7	-	+
pant388	4713459	4713664	4713100	4713795	206	PA4209	PA4210	<>>	359	131	-	+	+	-	-	-	+
pant389	4783131	4783217	4783112	4783227	87	PA4274	PA4275	<?<	19	10	-	-	+	-	-	-	+
pant390	4791760	4791860	4791725	4791950	101	PA4280.2	PA4280.3	<><	35	90	-	-	+	+	68.3	-	++
pant391	4813801	4813991	4812602	4814212	191	PA4289	PA4291	><	1199	221	-	+	+	-	-	-	+
pant392	4814535	4814996	4814215	4815042	462	PA4290	PA4292	<>>	320	46	+	+	+	-	-	-	-
pant393	4830663	4830846	4830555	4830963	184	PA4305	PA4306	<<>	108	117	-	+	+	+	10.5	-	-
pant394	4834913	4835263	4834865	4835263	351	PA4308	PA4309	<>>	48	0	-	-	+	-	-	-	-
pant395	4854991	4855070	4854979	4855071	80	PA4326	PA4327	<>>	12	1	-	+	-	+	14.8	-	-
pant396	4860508	4861088	4860256	4861652	581	PA4331	PA4333	<<>	252	564	-	+	+	+	2.6	-	+
pant397	4861124	4861399	4860256	4861652	276	PA4331	PA4333	<?>	868	253	+	+	+	-	-	-	-
pant398	4861733	4862026	4861712	4863272	294	PA4332	PA4334	<?<	21	1246	-	+	+	-	-	-	++
pant399	4862054	4862348	4861712	4863272	295	PA4332	PA4334	<><	342	924	-	+	+	+	0.4	-	++
pant400	4862626	4863181	4861712	4863272	556	PA4332	PA4334	<><	914	91	+	+	+	-	-	-	++
pant401	4878593	4878723	4878588	4878789	131	PA4349	PA4350	<<>	5	66	-	+	+	+	6.4	-	-
pant402	4888276	4888582	4888196	4889111	307	PA4360	PA4361	<><	80	529	+	+	+	+	0.4	-	+
pant403	4893521	4893633	4893351	4893696	113	PA4365	PA4366	>?>	170	63	-	-	+	-	-	-	-
pant404	4903327	4903450	4903297	4903465	124	PA4373	PA4374	>>>	30	15	-	+	-	+	0.4	-	-
pant405	4921096	4921416	4920867	4921980	321	PA4390	PA4392	<<>	229	564	-	+	+	-	-	-	+
pant406	4956222	4956326	4956030	4956327	105	PA4421	PA4421.1	<<>	192	1	-	+	+	-	-	-	-
pant407	4989152	4989259	4988892	4989304	108	PA4456	PA4457	<<>	260	45	-	-	+	+	5.7	-	-
pant408	5013486	5013605	5013432	5013670	120	PA4481	PA4482	<<>	54	65	-	-	+	-	-	-	+
pant409	5030276	5030516	5030249	5030524	241	PA4494	PA4495	<>>	27	8	-	+	+	-	-	-	-
pant410	5030390	5030516	5030249	5030524	127	PA4494	PA4495	<<>	141	8	-	+	+	-	-	-	-
pant411	5031356	5031482	5031236	5031487	127	PA4495	PA4496	>>>	120	5	-	+	-	-	-	-	-
pant412	5038743	5038856	5038676	5038900	114	PA4500	PA4501	>>>	67	44	-	-	+	+	6.6	-	-
pant413	5055878	5056088	5055878	5056094	211	PA4514	PA4515	<<>	0	6	-	+	-	-	-	-	-
pant414	5061315	5061556	5061279	5061729	242	PA4519	PA4520	<?<	36	173	-	+	+	+	0.2	-	-
pant415	5080300	5080649	5080209	5080756	350	PA4539	PA4540	<<>	91	107	-	+	+	-	-	-	-
pant416	5086761	5086889	5086697	5086924	129	PA4541	PA4541.1	>?<	64	35	-	+	-	+	0.3	-	-
pant417	5091879	5091950	5091816	5091960	72	PA4544	PA4545	<>>	63	10	-	+	-	-	-	-	-
pant418	5097463	5097563	5097463	5097596	101	PA4548	PA4549	<<>	0	33	-	+	+	-	-	-	-
pant419	5112382	5112481	5112380	5112662	100	PA4562	PA4563	<>>	2	181	-	+	-	-	-	-	-
pant420	5118198	5118323	5118196	5118537	126	PA4570	PA4571	>>>	2	214	-	+	+	-	-	-	-
pant421	5141281	5141583	5141190	5141785	303	PA4590	PA4591	<<<	91	202	-	+	+	-	-	-	-
pant422	5146428	5146671	5146404	5146906	244	PA4594	PA4595	<?<	24	235	-	+	+	+	3.6	-	-
pant423	5166749	5166864	5166698	5166899	116	PA4607	PA4608	<>>	51	35	-	+	+	+	4.3	-	-
pant424	5171891	5172157	5171727	5172252	267	PA4613	PA4614	<<>	164	95	-	+	+	+	3.7	-	-
pant425	5174710	5174954	5174653	5174979	245	PA4616	PA4617	<?>	57	25	+	-	-	-	-	-	-
pant426	5182285	5182374	5182269	5182578	90	PA4621	PA4622	<?>	16	204	-	-	+	-	-	-	-
pant427	5203330	5203607	5203243	5203667	278	PA4634	PA4635	<?>	87	60	+	+	+	+	0.2	-	-

pant428	5207886	5208410	5207623	5208463	525	PA4639	PA4640	<>>	263	53	+	+	+	-	-	-	-
pant429	5222265	5222539	5222261	5222539	275	PA4654	PA4655	<><	4	0	-	+	+	-	-	-	-
pant430	5224560	5224883	5224477	5224889	324	PA4656	PA4657	<?>	83	6	-	+	+	-	-	-	-
pant431	5242663	5242905	5242560	5243177	243	PA4674	PA4675	<<>	103	272	-	+	+	-	-	-	-
pant432	5242933	5243127	5242560	5243177	195	PA4674	PA4675	<>>	373	50	-	+	+	-	-	-	-
pant433	5247427	5247538	5247424	5247621	112	PA4677	PA4678	<<<	3	83	-	+	-	-	-	-	-
pant434	5248778	5248977	5248773	5249031	200	PA4679	PA4680	<<<	5	54	-	+	+	-	-	-	-
pant435	5269263	5269463	5269260	5269802	201	PA4690.5	PA4691	<?<	3	339	-	+	+	-	-	-	+
pant436	5269486	5269748	5269260	5269802	263	PA4690.5	PA4691	<?<	226	54	-	+	+	+	0.1	-	-
pant437	5275788	5275989	5275733	5276282	202	PA4696	PA4697	<<>	55	293	-	+	-	-	-	-	+
pant438	5297644	5297893	5297484	5297896	250	PA4718	PA4719	<>>	160	3	-	+	+	-	-	-	-
pant439	5309038	5309310	5308965	5309325	273	PA4726.2	PA4727	<>>	73	15	+	+	+	-	-	-	-
pant440	5323178	5323345	5323102	5323373	168	PA4739	PA4740	<><	76	28	-	+	-	+	0.3	-	-
pant441	5353511	5353695	5353508	5353782	185	PA4767	PA4768	<<>	3	87	-	+	+	-	-	-	-
pant442	5355250	5355360	5355059	5355386	111	PA4769	PA4770	<>>	191	26	-	+	+	+	0.3	-	-
pant443	5357104	5357204	5357076	5357229	101	PA4770	PA4771	><>	28	25	-	-	+	-	-	-	-
pant444	5361261	5361475	5361259	5361585	215	PA4772	PA4773	>>>	2	110	-	+	+	+	0.2	-	+
pant445	5372502	5372590	5372268	5372590	89	PA4784	PA4785	>><	234	0	-	+	+	-	-	-	-
pant446	5383814	5383886	5383813	5383946	73	PA4797	PA4798	<?>	1	60	-	+	-	-	-	-	-
pant447	5391673	5391785	5391555	5391835	113	PA4805	PA4806	<>>	118	50	-	+	+	-	-	-	-
pant448	5418130	5418343	5418125	5418350	214	PA4825	PA4826	<<<	5	7	-	+	+	+	0.2	-	-
pant449	5418593	5418778	5418570	5418898	186	PA4826	PA4827	<<<	23	120	+	+	-	+	0.3	-	-
pant450	5447495	5447604	5447445	5447647	110	PA4851	PA4852	>?>	50	43	-	+	+	-	-	-	-
pant451	5472178	5472292	5472043	5472438	115	PA4874	PA4875	<?>	135	146	-	+	-	+	3.4	-	-
pant452	5483773	5483927	5483769	5483987	155	PA4887	PA4888	<<<	4	60	-	+	+	-	-	-	-
pant453	5513435	5513583	5513113	5513731	149	PA4913	PA4914	<>>	322	148	-	+	-	+	16.7	-	-
pant454	5522145	5522246	5522111	5522385	102	PA4922	PA4923	<<>	34	139	-	-	+	-	-	-	-
pant455	5563373	5563543	5563288	5563964	171	PA4957	PA4958	>?>	85	421	-	+	+	+	2.5	-	-
pant456	5579547	5579621	5579500	5579652	75	PA4969	PA4970	<?<	47	31	-	+	-	-	-	-	-
pant457	5580773	5580854	5580720	5580961	82	PA4971	PA4972	<<>	53	107	-	-	+	-	-	-	-
pant458	5583649	5583870	5583646	5584100	222	PA4973	PA4974	<?>	3	230	+	+	+	-	-	TPP	-
pant459	5593485	5593669	5593425	5593682	185	PA4980	PA4981	<><	60	13	+	+	+	-	-	-	-
pant460	5647596	5647784	5647396	5647813	189	PA5020	PA5021	>?>	200	29	-	+	+	-	-	-	-
pant461	5654418	5654798	5654406	5654810	381	PA5023	PA5024	>?>	12	12	+	+	+	-	-	-	+
pant462	5689838	5689938	5689716	5689963	101	PA5050	PA5051	>?>	122	25	-	+	+	-	-	-	-
pant463	5695214	5695341	5694948	5695365	128	PA5055	PA5056	>>>	266	24	-	-	+	+	14.7	-	-
pant464	5714033	5714242	5714033	5714269	210	PA5076	PA5077	<<<	0	27	-	+	-	-	-	-	-
pant465	5718436	5718578	5718426	5718881	143	PA5078	PA5079	<<<	10	303	-	+	-	-	-	-	-
pant466	5718584	5718804	5718426	5718881	221	PA5078	PA5079	<?<	158	77	-	+	+	-	-	-	-
pant467	5766271	5766371	5766148	5766483	101	PA5118	PA5119	<>>	123	112	-	+	-	+	2.4	-	-
pant468	5775535	5775702	5774808	5776085	168	PA5125	PA5127	><>	727	383	-	+	+	+	3.6	-	-
pant469	5776459	5776595	5776088	5776605	137	PA5126	PA5128	<?<	371	10	-	+	+	+	0.3	-	+
pant470	5777944	5778124	5777810	5778133	181	PA5130	PA5131	<>>	134	9	-	-	+	+	0.2	-	-
pant471	5790193	5790397	5790133	5790444	205	PA5141	PA5142	<?<	60	47	-	+	+	-	-	-	+
pant472	5803741	5803812	5803514	5803866	72	PA5155	PA5156	>?>	227	54	-	+	+	-	-	-	-
pant473	5804955	5805202	5804803	5805205	248	PA5156	PA5157	>>>	152	3	-	+	+	-	-	-	-
pant474	5817046	5817146	5816951	5817190	101	PA5166	PA5167	>>>	95	44	-	-	+	+	5.0	-	-
pant475	5820448	5820845	5820115	5820909	398	PA5169	PA5170	>?>	333	64	+	+	+	+	5.9	-	-
pant476	5825725	5826053	5825720	5826134	329	PA5173	PA5174	><>	5	81	-	-	+	-	-	-	-
pant477	5835909	5835984	5835896	5835993	76	PA5182	PA5183	><>	13	9	-	-	+	-	-	-	-
pant478	5836444	5836841	5836402	5836909	398	PA5183	PA5184	><>	42	68	+	+	+	-	-	-	-
pant479	5843826	5843965	5843801	5844053	140	PA5190	PA5191	<<>	25	88	-	+	+	-	-	-	-
pant480	5853463	5853562	5853397	5853584	100	PA5200	PA5201	<<>	66	22	-	+	-	+	2.2	-	-
pant481	5859468	5859660	5859457	5859792	193	PA5204	PA5205	<<>	11	132	-	+	-	-	-	-	-

pant482	5892668	5892830	5892661	5892909	163	PA5234	PA5235	><	7	79	-	+	+	+	0.2	-	-
pant483	5906521	5906609	5906462	5906625	89	PA5244	PA5245	><	59	16	-	+	+	-	-	-	-
pant484	5914527	5914996	5914522	5915042	470	PA5252	PA5253	<>	5	46	+	+	+	-	-	-	-
pant485	5924373	5924582	5924372	5924595	210	PA5262	PA5263	<?>	1	13	-	+	+	-	-	-	-
pant486	5955066	5955166	5955056	5955179	101	PA5290	PA5291	><	10	13	-	+	-	+	2.3	-	-
pant487	5973554	5973827	5973477	5973832	274	PA5304	PA5305	<<<	77	5	+	+	+	-	-	-	-
pant488	5986290	5986454	5986171	5986474	165	PA5316.1	PA5317	><	119	20	-	+	+	-	-	-	-
pant489	5991138	5991429	5991132	5992382	292	PA5321	PA5322	>>>	6	953	-	+	+	-	-	-	-
pant490	5991564	5991777	5991132	5992382	214	PA5321	PA5322	>?>	432	605	-	+	+	-	-	-	+
pant491	5991800	5992035	5991132	5992382	236	PA5321	PA5322	>>>	668	347	-	+	+	-	-	-	-
pant492	5995818	5995936	5995812	5996035	119	PA5324	PA5325	<?>	6	99	-	+	+	-	-	-	-
pant493	6032310	6032385	6032269	6032386	76	PA5364	PA5365	<?<	41	1	-	+	+	-	-	-	-
pant494	6037875	6038041	6037790	6038211	167	PA5368	PA5369	<<<	85	170	-	+	+	-	-	-	+
pant495	6042772	6042872	6042737	6042962	101	PA5369.2	PA5369.3	<>	35	90	-	-	+	+	35.1	-	++
pant496	6045008	6045246	6044744	6045309	239	PA5369.5	PA5370	<<<	264	63	-	+	+	+	0.1	-	-
pant497	6051319	6051484	6051312	6051643	166	PA5374	PA5375	<?>	7	159	-	+	-	+	2.0	-	-
pant498	6056431	6056684	6056414	6056876	254	PA5378	PA5379	<>	17	192	-	+	+	-	-	-	-
pant499	6058343	6058460	6058254	6058672	118	PA5379	PA5380	>>>	89	212	-	+	+	-	-	-	-
pant500	6086677	6086895	6086675	6087098	219	PA5410	PA5411	<>	2	203	-	+	+	+	57.6	-	-
pant501	6125211	6125386	6125081	6125795	176	PA5440	PA5441	><	130	409	-	-	+	-	-	-	++
pant502	6125568	6125748	6125081	6125795	181	PA5440	PA5441	><	487	47	-	+	-	-	-	-	+
pant503	6162085	6162220	6161868	6162220	136	PA5472	PA5473	<>	217	0	-	+	+	-	-	-	-
pant504	6167543	6167752	6167410	6167782	210	PA5476	PA5477	<>	133	30	-	+	+	-	-	-	-
pant505	6172877	6172977	6172874	6173347	101	PA5482	PA5483	<?>	3	370	-	+	+	-	-	-	-
pant506	6173073	6173287	6172874	6173347	215	PA5482	PA5483	<>	199	60	-	+	+	-	-	-	-
pant507	6187926	6188150	6187878	6188165	225	PA5495	PA5496	>?<	48	15	-	+	+	-	-	-	-
pant508	6211770	6212005	6211769	6212146	236	PA5518	PA5519	<<<	1	141	-	+	+	-	-	-	-
pant509	6220755	6220874	6220740	6221100	120	PA5528	PA5529	<>	15	226	-	+	+	-	-	-	-
pant510	6240638	6240792	6240633	6240869	155	PA5546	PA5547	<?>	5	77	-	+	-	+	2.9	-	-
pant511	6241760	6241846	6241494	6241850	87	PA5547	PA5548	><	266	4	-	+	+	-	-	-	-
pant512	6259392	6259552	6259198	6259670	161	PA5565	PA5566	<>	194	118	-	+	+	-	-	-	-
pant513	6263613	6263715	6263565	6263804	103	PA5568	PA5569	<>	48	89	-	+	+	-	-	-	-

a The coordinates of the transcript predicted by RNA sequencing data.
b Length of the sRNA candidate predicted by RNA sequencing data.
c The middle arrow indicates the orientation of the sRNA, while the flanking arrows indicate the orientation of the adjacent genes. An unknown sRNA orientation is indicated by a question mark.
d A plus symbol (+) indicates that the sRNA is differentially expressed between the conditions tested in this work.
e The fold change is calculated as the ratio of the intensities of the sRNA between the early-stationary and exponential conditions. Thus, a fold change > 1 indicates that the sRNA is more highly expressed in the early-stationary than the exponential phase.
f Candidate sRNAs were queried against the Rfam database and matches to known sRNAs are indicated.
g The sequence conservation of candidate sRNAs in other microbial organisms was investigated using the BLASTN algorithm. (-), no sequence conservation found; (+), sequence conserved primarily in Pseudomonadaceae; (++) sequence conserved in bacterial species outside the Pseudomonadaceae family.

Supplementary Table 5. 5´-RACE validation of sRNA candidates.

Name	Start predicted by RNA sequencing[a]	Start validated by 5´-RACE[b]	Nt difference between RNAseq and 5´-RACE[c]	Direction[d]
pant44	363166	363154	12	<<<
pant66	648847	648847	0	<<<
pant217	2627303	2627304	1	<>>
RgsA (forward)	3318747	3318737	10	>><
RgsA (reverse)	3318868	3318870	2	>><
pant381	4629347	4629347	0	<<>
pant441	5353695	5353695	0	<<>
pant503	6162085	6162085	0	<>>

[a] Transcript start predicted by our script based on the RNA sequencing data.
[b] Experimentally found values using 5´-RACE.
[c] Difference in nucleotides between the start of transcription predicted by RNA sequencing and that determined by 5´-RACE.
[d] The middle arrow indicates the orientation of the sRNA, while the flanking arrows indicate the orientation of the adjacent genes.

Supplementary Table 6. Novel sRNA candidates conserved in organisms outside the Pseudomonadaceae family.

Name	Start[a]	End[a]	Intergenic region start	Intergenic region end	Length[b]	5' flanking gene	3' flanking gene	Direction[c]	Detection in Library A	Detection in Library B	Detection in Library C	Differential expression[d]	Fold change[e]	Rfam[f]	Orders[g]	Classes[h]
pant42	354489	354679	354462	354753	191	PA0314	PA0315	<?>	-	-	+	+	2.6	-	Burkholderiales	Beta Proteobacteria
pant57	521960	522172	521316	522464	213	PA0460	PA0462	><	+	+	+	-	-	-	Aeromonadales / Alteromonadales / Burkholderiales / unclassified Deltaproteobacteria / Enterobacteriales / Oceanospirillales / Pseudomonadales / unclassified Gammaproteobacteria (gamma proteobacterium HdN1) / Xanthomonadales	Beta Proteobacteria / Delta Proteobacteria / Gamma Proteobacteria
pant59	522753	522868	522518	523253	116	PA0461	PA0463	<>	-	-	+	+	8.4	-	Halomonas / Pseudomonas	Gamma Proteobacteria
pant95	927232	927551	927215	927879	320	PA0847	PA0849	<?>	-	-	+	+	6.9	-	Rhizobiales / Desulfovibrionales / Rhodocyclales / Gallionellales / Nitrosomonadales / Neisseriales / Methylophilales / Oceanospirillales / Alteromonadales / Vibrionales / Aeromonadales / Enterobacteriales / Methylococcales / Pseudomonadales	Alpha Proteobacteria / Beta Proteobacteria / Delta Proteobacteria / Gamma Proteobacteria
pant96	927603	927784	927215	927879	182	PA0847	PA0849	<>	-	+	+	-	-	-	Aeromonadales / Enterobacteriales / Neisseriales / Rhodocyclales	Beta Proteobacteria / Gamma Proteobacteria
pant120	1204789	1205025	1204782	1205030	237	PA1112	PA1112.1	><	+	+	+	+	3.0	-	Aeromonadales / Rhodocyclales / Pseudomonadales	Beta Proteobacteria / Gamma Proteobacteria
pant138	1372693	1372804	1372683	1372978	112	PA1263	PA1264	>?<	-	-	+	-	-	-	Aeromonadales / Burkholderiales / Xanthomonadales	Beta Proteobacteria / Gamma Proteobacteria
pant144	1471942	1472164	1471673	1472546	223	PA1359	PA1361	<><	-	+	+	-	-	-	Burkholderiales / Caulobacterales / Enterobacteriales / Myxococcales / Oceanospirillales / Pseudomonadales / Rhizobiales / Sphingomonadales / Xanthomonadales	Alpha Proteobacteria / Beta Proteobacteria / Delta Proteobacteria / Gamma Proteobacteria
pant160	1717940	1718216	1717895	1718385	277	PA1578	PA1579	>>>	+	+	+	+	0.2	-	Burkholderiales / Pseudomonadales	Beta Proteobacteria / Gamma Proteobacteria
pant184	2104152	2104333	2104097	2104596	182	PA1925	PA1926	>>>	-	+	+	-	-	-	Alteromonadales / Burkholderiales / Caulobacterales / Enterobacteriales / Methylophilales / Myxococcales / Neisseriales / Nitrosomonadales / Oceanospirillales / Pseudomonadales / Rhizobiales / Rhodocyclales / Sphingomonadales / Xanthomonadales	Alpha Proteobacteria / Beta Proteobacteria / Delta Proteobacteria / Gamma Proteobacteria
pant185	2104397	2104580	2104097	2104596	184	PA1925	PA1926	>>>	-	+	+	+	0.3	-	Alteromonadales / Burkholderiales / Methylophilales / Neisseriales / Oceanospirillales / Pseudomonadales / Rhodocyclales	Beta Proteobacteria / Gamma Proteobacteria
pant242	3037237	3037426	3035703	3037885	190	PA2684	PA2686	<<>	-	+	+	-	-	-	Burkholderiales / Pseudomonas / Rhodocyclales	Beta Proteobacteria / Gamma Proteobacteria
pant251	3094233	3094534	3094186	3094756	302	PA2733	PA2734	<?<	-	+	+	-	-	-	Acidithiobacillales / Desulfovibrionales / Nitrosomonas / Desulfuromonadales / Chlorobiales / Pseudomonadales / unclassified Gammaproteobacteria (gamma proteobacterium HdN1) / Synergistales	Chlorobea / Synergistia / Beta Proteobacteria / Delta Proteobacteria / Gamma Proteobacteria
pant277	3305612	3305787	3305565	3305918	176	PA2945	PA2946	<<>	-	+	+	-	-	obalami	Rhodobacterales / Pseudomonas / Rhizobiales	Alpha Proteobacteria / Gamma Proteobacteria
pant366	4472724	4472848	4472199	4473622	125	PA3991	PA3993	<<>	-	-	+	+	5.2	-	Burkholderiales / Caulobacterales / Pseudomonadales / Rhizobiales / Rhodobacterales / Rhodospirillales / unclassified Alphaproteobacteria (Polymorphum gilvum) / Xanthomonadales	Alpha Proteobacteria / Beta Proteobacteria / Gamma Proteobacteria
pant367	4472877	4473081	4472199	4473622	205	PA3991	PA3993	<>>	-	+	+	-	-	-	Pseudomonadaceae / Burkholderiales / Rhodospirillales / unclassified Alphaproteobacteria (Polymorphum gilvum)	Alpha Proteobacteria / Beta Proteobacteria / Gamma Proteobacteria
pant390	4791760	4791860	4791725	4791950	101	PA4280.2	PA4280.3	<><	-	-	+	+	68.3	-	unclassified Bacteria (toluene-degrading bacterium UCR 026m; chlamydia-associated clinical samples) / Pseudomonadales / Myxococcales	Delta Proteobacteria / Gamma Proteobacteria
pant398	4861733	4862026	4861712	4863272	294	PA4332	PA4334	<?<	-	+	+	-	-	-	Alteromonadales / Burkholderiales / Acidobacteriales / Chromatiales / Gallionellales / Oceanospirillales / Pseudomonadales / Rhodocyclales / Xanthomonadales	Acidobacteria / Beta Proteobacteria / Gamma Proteobacteria
pant399	4862054	4862348	4861712	4863272	295	PA4332	PA4334	<><	-	+	+	+	0.4	-	Aeromonadales / Burkholderiales / Chromatiales / Neisseriales / Pseudomonadales / Rhizobiales / Rhodocyclales / Sphingomonadales / unclassified Betaproteobacteria (Candidatus Accumulibacter) / Xanthomonadales	Alpha Proteobacteria / Beta Proteobacteria / Gamma Proteobacteria
pant400	4862626	4863181	4861712	4863272	556	PA4332	PA4334	<><	+	+	+	-	-	-	Burkholderiales / Chromatiales / Neisseriales / Oceanospirillales / Pseudomonadales / Rhodocyclales / Sphingomonadales / Xanthomonadales	Alpha Proteobacteria / Beta Proteobacteria / Gamma Proteobacteria
pant495	6042772	6042872	6042737	6042962	101	PA5369.2	PA5369.3	<><	-	-	+	+	35.1	-	unclassified Bacteria (toluene-degrading bacterium UCR 026m / chlamydia-associated clinical samples) / Pseudomonadales / Myxococcales	Delta Proteobacteria / Gamma Proteobacteria
pant501	6125211	6125386	6125081	6125795	176	PA5440	PA5441	>><	-	-	+	-	-	-	Burkholderia / Aeromonadales / Pseudomonas	Delta Proteobacteria / Gamma Proteobacteria

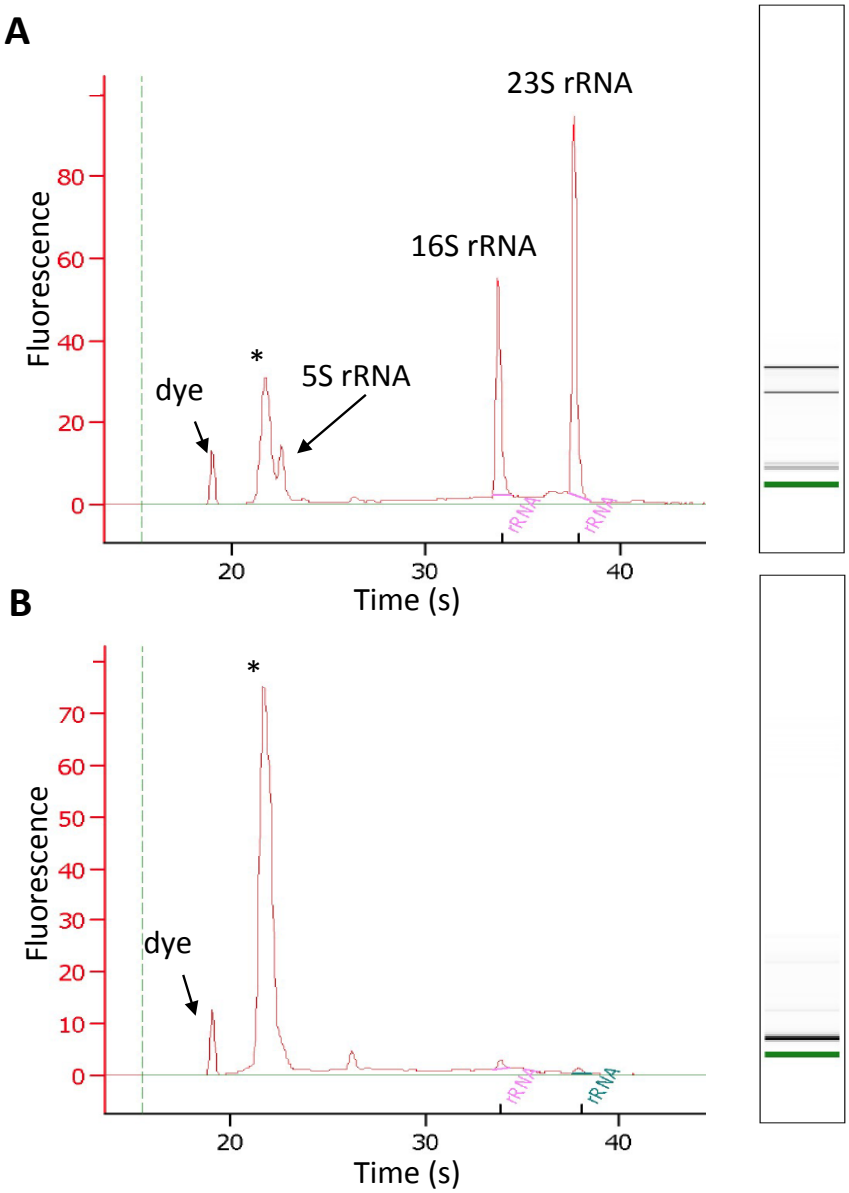
a Transcript's coordinates predicted by the RNA sequencing data.
b Length of the sRNA candidate predicted by the RNA sequencing data.
c The middle arrow indicates the orientation of the sRNA, while the flanking arrows indicate the orientation of the adjacent genes. Unknown orientation of sRNAs is indicated by a question mark.
d A + indicates that the sRNA is differentially expressed between the conditions tested in this work.
e The fold change indicates the ratio of the intensities of the sRNA between the early-stationary and exponential conditions. Thus, a number higher than 1 indicates that the sRNA is more expressed in the early-stationary than the exponential phase.
f Candidate sRNAs was queried against the Rfam database; matches to known sRNAs are indicated.
g Orders of the bacteria in which the sequence of the sRNA candidate was found to be conserved.
h Classes of the bacteria in which the sequence of the sRNA candidate was found to be conserved.

Supplementary Table 7. Cut-offs used for the detection of novel transcripts.

Library	Coverage at each base	Average expression
LIB>100	>20	>40
LIB<500	>35	>70
LIB<200	>50	>100

Supplementary Figure 1. Agilent Bioanalyzer analysis of total and rRNA-depleted RNA.

The 23S, 16S and 5S rRNAs were removed from the total RNA (A) by subtractive hybridization as described in the Experimental Procedures. Note that the peaks corresponding to 23S, 16S and 5S rRNAs are not present in the rRNA-depleted sample (B). The peaks marked with asterisks (*) probably correspond to tRNAs and sRNA species shorter than 120 nt.



Article 2

Identification of bacterial small RNAs by RNA sequencing

Gómez-Lozano, M., Marvig, R. L., Molin, S., & Long, K. S.
Methods in Pseudomonas aeruginosa, **2013**. *Accepted for publication.*

Identification of bacterial small RNAs by RNA sequencing

María Gómez-Lozano[1,2], Rasmus Lykke Marvig[1], Søren Molin[1,2], and Katherine S. Long[1,2,*]

[1]Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.
[2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.

*To whom correspondence should be addressed: e-mail: kalon@biosustain.dtu.dk

ABSTRACT

Small regulatory RNAs (sRNAs) in bacteria are known to modulate gene expression and control a variety of processes including metabolic reactions, stress responses and pathogenesis in response to environmental signals. A method to identify bacterial sRNAs on a genome-wide scale based on RNA sequencing (RNA-seq) is described that involves the preparation and analysis of three different sequencing libraries. As a significant number of unique sRNAs are identified in each library, the libraries can either be used alone or in combination to increase the number of sRNAs identified. The approach may be applied to identify sRNAs in any bacterium under different growth and stress conditions.

1. INTRODUCTION

Small RNAs (sRNAs) in bacteria play important regulatory roles in controlling various physiological processes such as carbon metabolism, iron homeostasis, and virulence in response to environmental cues. Bacteria contain hundreds of sRNAs that exist as a heterogeneous group of transcripts with a typical size range of 70 to 500 nucleotides (nt) (1,2). Many sRNAs function by affecting the expression of mRNA targets via base-pairing, while others act by binding to proteins and altering their activity. The base-pairing sRNAs fall into two broad classes: *cis*-encoded antisense sRNAs and *trans*-encoded sRNAs. The former are encoded on the DNA strand opposite the target RNA and therefore have perfect complementarity with their targets. *Trans*-encoded sRNAs are encoded in intergenic

regions and often have multiple targets with which there is only limited complementarity. Despite the fact that some bacterial sRNAs have been known for some time, earlier studies to identify these RNAs have often depended on serendipity, direct detection due to high abundance, protein copurification, and computational predictions based on sequence conservation (*3*). Recent efforts to identify sRNAs on a genome-wide scale have been based mainly on the extremely sensitive, probe-independent RNA sequencing approach (RNA-seq) that has revolutionized the field and led to the discovery of hundreds of novel sRNAs in diverse bacteria (4,5).

We have recently developed a robust RNA-seq method and used it to identify over 500 novel sRNAs in *P. aeruginosa* (*6*). The approach is based on three different sequencing libraries prepared from different RNA populations that can either be used alone or in combination to increase the number of identified sRNAs. In order to enrich the RNA samples for sRNA detection, ribosomal RNAs (23S, 16S and 5S rRNAs) are depleted from the total RNA. One of the libraries (LIB>100) is prepared using a standard RNA-seq protocol for full transcriptomes and contains information on all mRNAs transcribed by the bacteria. The other two libraries (LIB<500 and LIB<200) are prepared from size-selected RNA containing transcripts shorter than 500 and 200 nt, respectively. A significant number of unique sRNAs are detected in each library, suggesting that the set of identified RNAs obtained depends strongly on the specific library preparation strategy used (*6*). The detailed experimental procedures to prepare the three sequencing libraries are described in the first sections of the chapter.

Using RNA-seq allows counting the number of reads that align to specific parts of a genome, producing results similar to those of gene expression microarrays. While the end results are similar, the informatic challenges of analyzing RNA-seq data are fundamentally different than those of analyzing microarrays. Although there are several ways in which RNA-seq data can be analyzed, our analysis pipeline to find novel *P. aeruginosa* intergenic and antisense sRNAs is detailed in the latter sections of this chapter.

2. MATERIALS (see Note 1)

2.1 General materials and equipment

1. Disposable, nuclease-free pipette tips with filter, 1.5-mL microcentrifuge tubes, 15-mL conical tubes.

2. RNase-ZAP (Ambion) to remove ribonuclease (RNase) contamination from glass and plastic (see Note 1).
3. Glycogen (5 mg/mL) (Ambion).
4. Nuclease-free water (not DEPC-treated water).
5. TE buffer (10 mM Tris-HCl, pH 7, 1 mM EDTA).
6. 100% and 70% Ethanol.
7. Phase-Lock Gel tubes, Heavy, 2 mL (5 PRIME).
8. 3 M Sodium acetate (NaOAc), pH 5.5 (Ambion).
9. PCI (Phenol:Chloroform:Isoamyl Alcohol) solution (25:24:1, v/v).
10. RiboLock RNase inhibitor (40 U/μL) (Thermo Scientific).
11. Microcentrifuges (room temperature and 4ºC), centrifuge (4ºC).
12. Vortexer.
13. Heat block.
14. Bioanalyzer (Agilent).
15. RNA 6000 Nano Kit (Agilent).
16. Qubit® Fluorometer (Invitrogen).
17. Qubit® dsDNA HS or BR Assay Kit.
18. Qubit® RNA Assay Kit.
19. Low-volume spectrophotometer, for example the Nanodrop (Thermo Fisher Scientific).

2.2 Preparation of total RNA

1. Isopropanol.
2. Trizol (Invitrogen).
3. Chloroform.
4. Phenol.
5. Deoxyribonuclease (DNase) I, nuclease-free (1 U/μL) (Thermo Scientific).

2.3 Depletion of ribosomal RNAs (rRNAs)

1. MICROB*Express*™ Bacterial mRNA Enrichment Kit (Ambion).
2. DynaMag-2 Magnet (Invitrogen).
3. HPLC-purified 5S oligo:
   5'-AAAAAAAAAAAAAAAAAAGCGTTTCACTTCTGAGTTCGGCA-3'

2.4 Size selection of transcripts shorter than 500 nt (for LIB<500 library)

1. Mini-PROTEAN Electrophoresis System for vertical gel electrophoresis (Bio-Rad).
2. 10% Mini-PROTEAN TBE-Urea Gel (Bio-Rad, cat no. 456-6033). This is a precast 10% polyacrylamide gel (10-well, 30 µL well capacity) for use with the Mini-PROTEAN system (see Note 2).
3. Gel Loading Buffer II: 95% Formamide, 18mM EDTA, 0.025% SDS, 0.025% Xylene Cyanol, 0.025% Bromophenol Blue (Ambion).
4. Nuclease-free, sterile 10X TBE solution: 890 mM Tris base, 890 mM boric acid, 20 mM EDTA, pH 8.0. From this stock prepare 1xTBE solution.
5. 0.4 M sodium chloride (NaCl) solution.
6. Low range ssRNA ladder (New England Biolabs).
7. 5 µm filter tube (IST Engineering).
8. Gel Breaker tube (IST Engineering).
9. Sterile scalpel blade.
10. SYBR Gold Nucleic Acid Gel Stain (Invitrogen).
11. Tube shaker or rotator.
12. Safe Imager 2.0 Blue Light Transilluminator (Invitrogen) or similar

2.5 Library-specific RNA preparation steps

1. *mir*Vana™ miRNA Isolation Kit (Ambion) (for LIB<200 library).
2. Tobacco Acid Pyrophosphatase (Epicentre) (for LIB<500 and LIB<200 libraries).
3. RNase III (Invitrogen) (for LIB<500 library).

2.6 Library preparation

2.6.1 Library preparation (all libraries)

1. Magnetic Stand-96 (Ambion).
2. 96-well thermal cycler.
3. 10-µL, 200-µL, and 1000-µL multichannel pipettes.
4. Agencourt AMPure XP 60 mL kit (Beckman Coulter Genomics).
5. 96-well 250-µL PCR plates and caps for PCR strips compatible with the plates.
6. 8-well PCR strip tubes and caps.
7. SuperScript II Reverse Transcriptase (Invitrogen).
8. 10 mM Tris-Cl, pH 8.5, solution containing 0.1% Tween 20.
9. Microplate centrifuge.

2.6.2 Library preparation (for LIB>100 library)

1. TruSeq™ RNA Sample Preparation Kit v2 (Illumina) (see Notes 3 and 4).
2. DNA 1000 kit (Agilent).

2.6.3 Library preparation (for LIB<500 and LIB<200 libraries)

1. TruSeq™ Small RNA Sample Preparation Kit (Illumina) (see Notes 3 and 5).
2. T4 RNA Ligase 2, truncated (New England Biolabs).
3. High Sensitivity DNA kit (Agilent).

2.7 List of programs used

The programs referred to in the data analysis pipeline are listed below with name, version, and URL:

1. Flexbar v2.2, http://sourceforge.net/projects/theflexibleadap/
2. Bowtie 2, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
3. SAMtools v1.4, http://samtools.sourceforge.net/
4. Tablet v1.12.08.29, http://bioinf.scri.ac.uk/tablet/

3. METHODS

A general scheme of the method is illustrated in Figure 1, and the library characteristics are summarized in Table 1. The libraries prepared in this protocol are compatible with the Illumina technology (see Note 6). The sequencing libraries are called LIB>100, LIB<500, and LIB<200 to indicate the transcript sizes in nucleotides that are not excluded from each, but all transcripts in these size ranges are not necessarily detected in the corresponding library.

Table 1. Library characteristics.

	Library LIB>100	**Library LIB<500**	**Library <200**
Length of transcripts sequenced	>100 nt	<500 nt	<200 nt
Fragmentation	Divalent cations	RNase III	No
Strand specificity	No	Yes	Yes

The three libraries were prepared with different RNA populations isolated from bacterial total RNA. Library LIB>100 contains information corresponding to all transcripts in the cell, with the exception of rRNAs and transcripts shorter than 100 nt. This library was prepared using a standard RNA-seq protocol and is suitable for analyzing full transcriptomes. Although many known sRNAs are roughly 100 nt in length, transcripts shorter than 100 nt are not represented in this library because the corresponding cDNA fragments are not retained during the protocol. Size-selected RNA shorter than 500 and 200 nt and depleted of 5S rRNA was used to prepare the LIB<500 and LIB<200 libraries, respectively. As these libraries were prepared from RNA that does not include longer transcripts, they are enriched for information on sRNAs and thus more sensitive for sRNA detection compared to the LIB>100 library that contains information on the full transcriptome (*6*). Fragments corresponding to transcripts shorter than 100 nt are retained in the LIB<500 and LIB<200 libraries because specific adapters are ligated to the 5´- and 3´-ends of RNAs prior to cDNA synthesis. In addition, they are strand-specific, which means that the information about the strand from which the transcripts are transcribed is retained. Furthermore, treatment of RNA used to prepare the LIB<500 and LIB<200 libraries ensures the inclusion of primary transcripts containing a 5´-triphosphate (see Note 7).

The library preparation protocols also differ in terms of whether they include fragmentation steps. In order to be compatible with the relatively short read-lengths of the Illumina sequencing technology, long RNA molecules must be cleaved into shorter fragments. Therefore, the protocols for preparation of the LIB>100 and LIB<500 libraries include fragmentation steps using divalent cations and RNase III, respectively. Library LIB<200 does not have any fragmentation-associated biases because the short RNAs used to prepare the library obviate the need for a fragmentation step.

In sections 3.12 to 3.17, we describe a high-throughput bioinformatics pipeline for the analysis of RNA-Seq. The objectives of the pipeline are to detect transcripts and to quantify their relative abundance between different samples. Our pipeline is designed for use on organisms where an annotated reference genome exists, and as input uses the quality-filtered and de-multiplexed sequencing reads obtained as output from sequencing centers (*i.e.* reads that are already in or can be converted to the FASTQ-format). Specifically, the pipeline was tested on 100-bp paired-end Illumina HiSeq2000 reads obtained from cDNA-libraries prepared from cultures of *Pseudomonas aeruginosa* PAO1 (*6*).

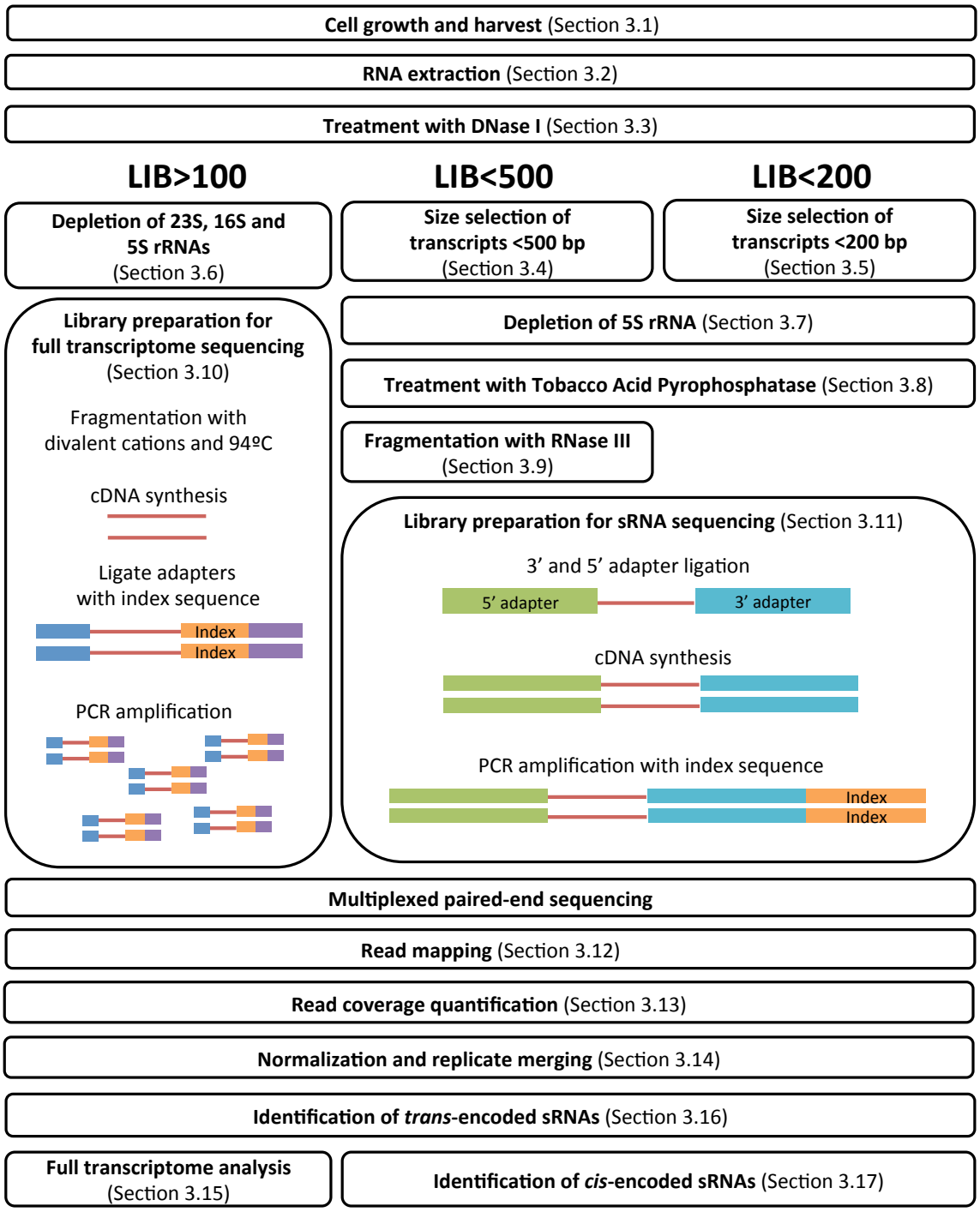


Figure 1. **Library preparation scheme.** Steps involved in the preparation of the different sequencing libraries (LIB>100, LIB<500 and LIB<200).

3.1 Cell Growth and Harvest

1. Prepare a 15-mL conical tube with 2 mL of the STOP solution (95% ethanol, 5% phenol). Pre-chill on ice.
2. Dilute an overnight bacterial culture in 50-100 volumes of growth medium.

3. When the desired OD_{600} is reached, transfer 10 mL of culture into the 15-mL conical tube with the STOP solution.
4. Vortex thoroughly for 15 s. Incubate at room temperature for 5 min.
5. Pellet the bacteria by centrifugation (3,500 x g, 10 min, 4ºC). Remove the supernatant.
6. Dissolve the pellet in 1 mL Trizol by pipetting (see Notes 8 and 9). Incubate at room temperature for 5 min.
7. Proceed to RNA extraction (section 3.2) or store at -80ºC. The homogenized samples can be stored at -80ºC for at least one month.

3.2 RNA extraction

1. Spin down the Phase-Lock Gel (PLG) at 12,000-16,000 x g for 30 s.
2. Transfer the homogenized sample to the PLG tube.
3. Add 200 µL of chloroform (work inside a fume hood). Shake vigorously for 30 s (do not vortex).
4. Incubate at room temperature for 2 min.
5. Centrifuge at 12,000-16,000 x g for 5 min at room temperature. The PLG will form a barrier between the aqueous and organic phases.
6. Transfer 500 µL of the aqueous phase to a clean 1.5-mL tube.
7. Add 400 µL isopropanol, mix and incubate for 10 min at room temperature.
8. Pellet the RNA by centrifugation (14,500 x g, 10 min, 4ºC). Remove the supernatant.
9. Wash the RNA pellet by adding 500 µL ice-cold 75% ethanol and vortex briefly.
10. Pellet the RNA by centrifugation (14,500 x g, 5 min, 4ºC). Remove the supernatant (see Note 10).
11. Repeat the last two steps for a total of two washes with ethanol. It is important to remove the ethanol completely at the end of the second wash.
12. Air-dry the pellet at room temperature for 5 min by leaving the tube lid open.
13. Dissolve the RNA pellet in 100 µL of nuclease-free water.
14. Incubate at 65ºC with shaking for 5 min.
15. Proceed to DNase I treatment (section 3.3) immediately or store at -80ºC.

3.3 Treatment with DNase I

1. Add to RNA sample (100 µL):
   1 µL RiboLock RNase inhibitor (40 U/µL)
   40 µL 10X DNase I reaction buffer with $MgCl_2$
   219 µL nuclease-free water

40 µL DNase I 1U/µL

400 µL total volume

2. Mix by gently pipetting the entire volume ten times.
3. Incubate at 37ºC for 30 min.
4. Spin down the Phase-Lock Gel (PLG) at 12,000-16,000 x g for 30 s.
5. Transfer the reaction mixture to the PLG tube.
6. Add 400 µL PCI solution.
7. Shake vigorously for 30 s (do not vortex).
8. Centrifuge at 12,000-16,000 x g for 5 min at room temperature.
9. Transfer the aqueous phase to a clean 1.5-mL tube (volume should be approximately 400 µL).
10. Precipitate the RNA by adding 0.02 volume glycogen (5 mg/mL), 0.1 volume 3 M NaOAc pH 5.5, and 2.5 volumes of 100% ethanol (ice-cold). Precipitate at -20ºC for at least one hour. At this point, the RNA can be stored at -20ºC or -80ºC (see Note 10). Pellet the RNA by centrifugation (14,500 x g, 30 min, 4ºC) and remove the supernatant. Recover the RNA as described in steps 9-12 of section 3.2.
11. Resuspend in 50 µL TE buffer.
12. Check the concentration and purity of the RNA sample using a low volume spectrophotometer (see Note 11). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer (see Note 12). The RNA profile should appear similar to that in Figure 2A. Proceed to the next step immediately (see Figure 1) or store at -80ºC.

3.4 Size selection of transcripts shorter than 500 nt (LIB<500 library)

Denaturing polyacrylamide-urea gels are used in order to size-select transcripts shorter than 500 nt (see Note 2).

1. Prior to sample loading the gels are pre-run in 1X TBE at 300V for 30 minutes.
2. Combine on ice total RNA (up to 150 µg) in a maximum volume of 30 µL with one volume of Gel Loading Buffer II. Each sample is loaded onto two lanes of the gel.
3. For each lane containing the marker, combine on ice:

   6 µL of Low Range ssRNA Ladder (3 µg)

   4 µL nuclease-free water

   10 µL Gel Loading Buffer II

   20 µL total volume
4. Incubate the RNA and marker at 65ºC for 5 minutes just prior to loading them onto the gel.

5. Clean the wells of the gel with a syringe containing 1X TBE buffer.
6. For each RNA sample load equal volumes into two adjacent wells on the gel. Leave an empty lane next to the RNA ladder and also between different samples.
7. Run the gel at 200 V for 35 minutes.
8. Stain the gel by soaking it in 50 mL 1X TBE with 5 µL of SYBR Gold Nucleic Acid Gel Stain at room temperature for 30 min. Use a nuclease-free container that is protected from light.
9. Place the 0.5-mL Gel Breaker tube into a 2-mL microcentrifuge tube.
10. Visualize the RNA and cut out the gel containing RNA between 50 and 500 nt by comparison with the bands of the RNA ladder.
11. Place the gel pieces corresponding to the same sample into the 0.5-mL Gel Breaker tube.
12. Centrifuge the stacked tubes (20,000 x g, 2 min, room temperature) to move the gel through the holes and into the 2-mL tube. Ensure that all of the gel has moved into the bottom tube.
13. Add 400 µL 0.4 M NaCl solution to the gel debris.
14. Elute the RNA by rotating or shaking the tube overnight at room temperature.
15. Transfer the eluate and the gel debris to the top of a 5-µm filter tube.
16. Centrifuge the filter at 600 x g for 10 s. Discard the filter and retain the eluate which contains the RNA.
17. Precipitate the RNA by following step 10 under section *3.3 Treatment with DNase I*.
18. Resuspend in 20 µL TE buffer.
19. Quantitate the yield of the rRNA-depleted RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen), or a low-volume spectrophotometer (see Note 11). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer (see Note 12). The profile of the RNA obtained should be similar to that depicted in Figure 2C.
20. Proceed to the 5S rRNA depletion (section 3.7) immediately or store at -80ºC.

3.5 Size selection of transcripts shorter than 200 nt (LIB<200 library)

The *mir*Vana™ miRNA Isolation Kit.from Ambion (Part Number 1560M Rev. C January 2011 / http://tools.invitrogen.com/content/sfs/manuals/fm_1560.pdf) can be used to select for sRNAs shorter than 200 nt (see Note 13).

1. Follow the protocol described in section *IV. Additional Procedures. A. Isolation of Small RNAs from Total RNA Samples* of the kit manual to obtain a 100 µL solution containing the sRNAs ≤ 200 nt.
2. Precipitate the RNA by following step 10 from section *3.3 Treatment with DNase I*.

3. Resuspend the RNA in 20 µL TE buffer.
4. Quantitate the yield of the rRNA-depleted RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen), or a low-volume spectrophotometer (see Note 11). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer (see Note 12). The profile of the RNA obtained should be similar to that depicted in Figure 2C.
5. Proceed to the 5S rRNA depletion (section 3.7) immediately or store at -80ºC.

3.6 Depletion of 23S, 16S and 5S rRNAs (LIB>100 library)

After obtaining total RNA, the next step is to deplete abundant rRNAs from the sample. Bacterial total RNA consists of approximately 86% rRNA **(7)**. Removal of rRNA significantly increases the fraction of mRNAs and sRNAs in the sample, allowing higher sequencing coverage and more sensitive detection of novel transcripts. Although the MICROB*Express*™ Bacterial mRNA Enrichment Kit from Ambion is used here (Part Number 1905M Rev. C January 2011 / http://tools.invitrogen.com/content/sfs/manuals/fm_1905.pdf), there are other commercially available kits to deplete rRNA (see Note 14). The MICROB*Express*™ Kit is designed to enrich bacterial mRNA (including sRNAs) from purified total RNA by removing the 16S and 23S rRNAs via a subtractive hybridization approach.

The MICROB*Express*™ Kit is followed with the following two modifications: an additional oligo is employed to remove 5S rRNA and 25% more magnetic beads are used. The modifications needed to remove the 23S, 16S and 5S rRNAs in a single step are described below.

1. Using the concentration obtained by the spectrophotometer, calculate the volume of sample that contains 2-10 µg of RNA (see Note 15). The recommended maximum amount of RNA per reaction is 10 µg and the recommended maximum volume is 15 µL. If the recommended maximum amount of total bacterial RNA is exceeded, then rRNA removal will be incomplete.
2. Follow the protocol detailed in *Section B Anneal RNA and Capture Oligonucleotide Mix* of the kit manual.
3. After step B.2, add 1 µL of the HPLC-purified 5S oligo (20 µM). Continue the protocol from step B.3.
4. In step C.1, remove 62.5 µL of Oligo MagBeads for each RNA sample, corresponding to 25% more beads per sample. The reason for this is that adding the HPLC-purified 5S oligo results in a 25% increase in capture oligos. We have successfully processed

Oligo MagBeads for up to 10 samples (625 µL) in a single 1.5-mL tube. Continue the protocol from step C.2.

5. In step D.2, add 62.5 µL of the prepared Oligo MagBeads to the RNA/Capture Oligo Mix and incubate at 37ºC for 15 min. Continue the protocol from step D.3.
6. In step E.2, resuspend the RNA pellet in 25 µL of nuclease-free water. Do not resuspend in TE buffer. Continue the protocol from step E.3 (but do not perform step E.4, as this removes sRNAs).
7. Quantitate the yield of the rRNA-depleted RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer. The profile of the RNA obtained should be similar to that depicted in Figure 2D.
8. Proceed to the LIB>100 library preparation immediately (section 3.10) or store at -80ºC.

3.7 Depletion of 5S rRNA after size selection (LIB<500 and LIB<200 libraries)

Although the 23S and 16S rRNAs are depleted from the samples after the size selection step, a significant fraction of the sample now consists of 5S rRNA. In order to deplete 5S rRNA, the MICROB*Express*™ Kit manual (Part Number 1905M Rev. C January 2011 / http://tools.invitrogen.com/content/sfs/manuals/fm_1905.pdf) is followed with the modification that the HPLC-purified 5S oligo is added as the only Capture Oligo as described below.

1. Using the concentration obtained by the spectrophotometer, calculate the volume of sample that contains 2-10 µg of RNA (see Note 15). The recommended maximum amount of RNA per reaction is 10 µg and the recommended maximum volume is 15 µL. If the recommended maximum amount of total bacterial RNA is exceeded, then rRNA removal will be incomplete.
2. Follow the protocol detailed in *Section B Anneal RNA and Capture Oligonucleotide Mix* of the kit manual. In step B.2, add 4 µL of the HPLC-purified 5S oligo (20 µM), but do not add 4 µl of Capture Oligo Mix. Continue the protocol from step B.3. Do not add the Capture Oligo Mix included in the kit, as this contains oligos complementary to 23S and 16S rRNAs.
3. In step E.2, resuspend the RNA pellet in 25 µL of nuclease-free water. Do not resuspend in TE buffer. Continue the protocol from step E.3 (but do not perform step E.4, as this removes sRNAs).

4. Quantitate the yield of the rRNA-depleted RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer. The profile of the RNA obtained should be similar to that depicted in Figure 2D.
5. Proceed to the Tobacco Acid Pyrophosphatase treatment (section 3.8) immediately or store at -80ºC.

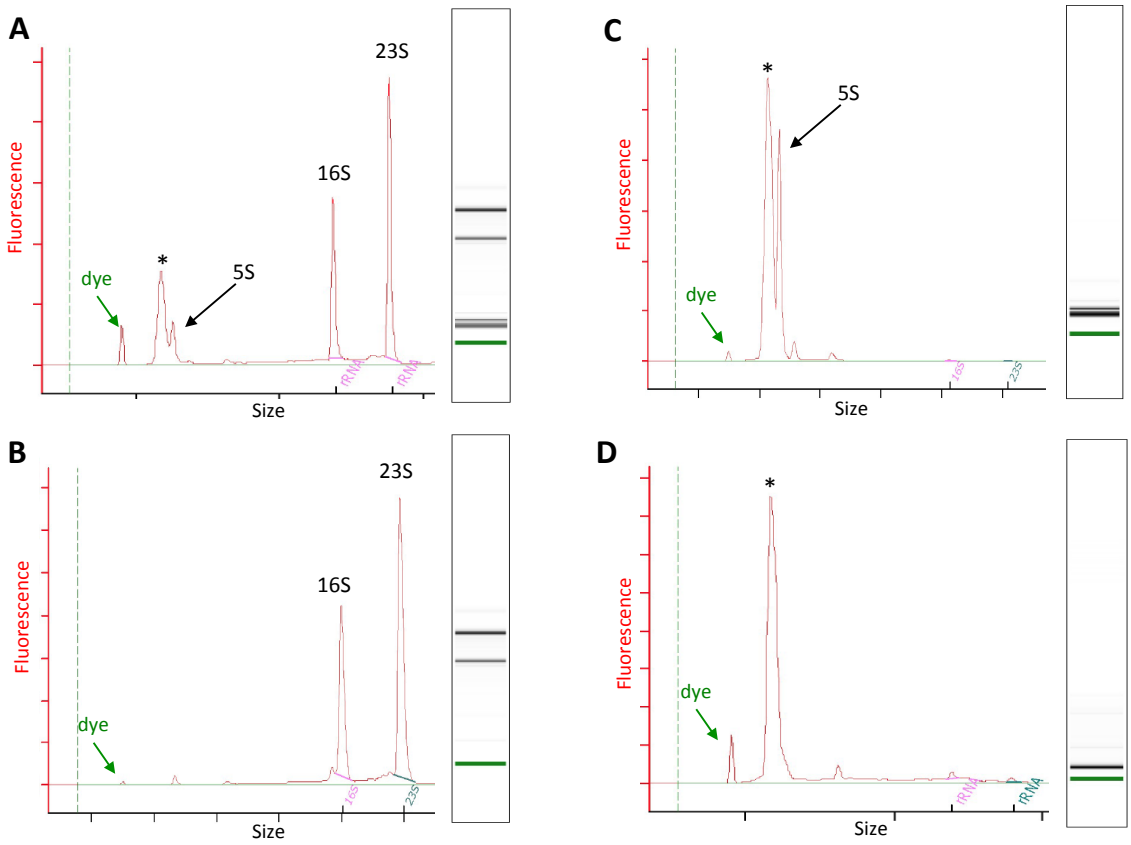


Figure 2. Typical size profiles of RNA samples. All profiles correspond to RNA extracted from *P. aeruginosa* PAO1 cells analyzed with the RNA 6000 Nano chip on the Bioanalyzer (Agilent). **A.** Total RNA extracted with Trizol and treated with DNase I. **B.** Total RNA extracted with the NucleoSpin RNA II kit (Macherey Nagel) and treated with DNase I. **C.** Size-selected RNA (<500 nt) purified on a denaturing polyacrylamide-urea gel. **D.** An rRNA-depleted sample. The 23S (2891 nt), 16S (1536 nt) and 5S (120 nt) rRNAs were removed from the total RNA by subtractive hybridization as described in the Materials and Methods sections. The peaks marked with asterisks (*) correspond to tRNAs and sRNA species shorter than 120 nt.

3.8 Treatment with Tobacco Acid Pyrophosphatase (LIB<500 and LIB<200 libraries)

After size-selection of RNA (used to prepare LIB<500 or LIB<200 libraries), the RNA is treated with Tobacco Acid Pyrophosphatase (TAP). TAP is used to convert 5′-triphosphate RNA into 5′-monophosphate RNA. This step is necessary to ensure that primary transcripts (containing a 5′-triphosphate) are ligated to adapters during the library preparation.

1. Assemble the reaction in a PCR tube on ice in the order given:
   17 µL Size-selected RNA (up to 10 µg) diluted in nuclease-free water
   2 µL 10X TAP Reaction Buffer
   1 µL TAP (10 U/ µL)
   20 µL total volume
2. Pipet up and down a few times to mix. Incubate at 37ºC for one hour.
3. Spin down the Phase-Lock Gel (PLG) at 12,000-16,000 x g for 30 s.
4. Transfer the sample to the PLG tube.
5. Add 130 µL nuclease-free water.
6. Add 150 µL PCI solution.
7. Shake vigorously for 30 s (do not vortex).
8. Centrifuge at 12,000-16,000 x g at room temperature for 5 min.
9. Transfer the aqueous phase to a clean 1.5-mL tube (the volume should be approximately 150 µL).
10. Precipitate the RNA by following step 10 from section *3.3 Treatment with DNase I*.
11. Resuspend in 15 µL nuclease-free water.
12. Quantitate the yield of the RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer. The profile of the RNA obtained should be similar to that depicted in Figure 2D.
13. Proceed to the next step immediately (see Figure 1) or store at -80ºC.

3.9 Fragmentation with RNase III (LIB<500 library)

During preparation of the LIB<500 library, the size-selected RNA needs to be fragmented after the TAP treatment. The fragmentation is performed with RNase III, which cleaves long double-stranded RNA (dsRNA) into shorter dsRNAs containing 5'-PO_4, and 3'-OH termini, and a 3' dinucleotide overhang. The fragmentation procedure should produce a distribution of RNA fragment sizes from 35 nt to several hundred nucleotides, where the average size is 100–200 nt.

1. Assemble the reaction in a microcentrifuge tube on ice in the order given:
   32 µL RNA (2 µg, from step 13, section 3.8) diluted in nuclease-free water
   4 µL 10X RNase III Buffer
   4 µL RNase III
   40 µL total volume
2. Pipet up and down a few times to mix. Incubate at 37ºC for 10 min.

3. Immediately after the incubation, add 110 µL of nuclease-free water and place the fragmented RNA on ice.
4. Spin down the Phase-Lock Gel (PLG) at 12,000-16,000 x g for 30 s.
5. Transfer the sample to a PLG tube.
6. Add 150 µL PCI solution.
7. Shake vigorously for 30 s (do not vortex).
8. Centrifuge at 12,000-16,000 x g for 5 min at room temperature.
9. Transfer the aqueous phase to a clean 1.5-mL tube (the volume should be approximately 150 µL).
10. Precipitate the RNA by following step 10 from section *3.3 Treatment with DNase I*.
11. Resuspend in 15 µL nuclease-free water.
12. Quantitate the yield of the fragmented RNA using the Qubit® RNA Assay Kit on the Qubit® Fluorometer (Invitrogen). Check the quality of the RNA with a RNA 6000 Nano chip on the Bioanalyzer. The profile of the RNA obtained should be similar to that depicted in Figure 2D. Proceed to the LIB<500 library preparation (section 3.11) immediately, or store at -80ºC.

3.10 Library Preparation for full transcriptome sequencing (LIB<100 library)

The LIB>100 library has been prepared with the Illumina TruSeq™ RNA Sample Preparation Kit v2 (Part Number 15026495 Rev. D September 2012 / http://support.illumina.com/documents/MyIllumina/b386d5c9-c919-48db-bdc1-8a687ba2a101/TruSeq_RNA_SamplePrep_v2_Guide_15026495_D.pdf) (see Notes 3 and 4) with some modifications that are explained below. Read the *Best Practices* section of the manual carefully. We perform the Low Sample (LS) protocol. If only 2-4 samples are to be pooled, read the *Pooling Guidelines* section of the manual.

1. Start the protocol with 100-400 ng of rRNA-depleted RNA, diluted in a maximum volume of 5 µL nuclease-free water (see Note 15).
2. Add 13 µL Elute, Prime, Fragment Mix to the sample.
3. Continue the protocol from step *Incubate RFP* (included in section *Purify and Fragment RNA*). Perform all the steps included in sections *Incubate RFP, Synthesize First Strand cDNA, Synthesize Second Strand cDNA, Perform End Repair, Adenylate 3' Ends, Ligate Adapters* and *Enrich DNA Fragments* (see Note 16). At this point the amplified libraries are in a 30 µL volume.
4. Quantify the libraries using the Qubit® dsDNA HS or BR Assay Kit on the Qubit® Fluorometer (Invitrogen) (see Note 17).

5. Check the size and purity of the sample by running a DNA 1000 chip on the Bioanalyzer. The final product should be a broad band with an average size of approximately 260 bp (see Figure 3A).
6. Adjust the concentration of the libraries to 10 nM using a 10 mM Tris-Cl (pH 8.5) solution containing 0.1% Tween 20.
7. Combine 10 µL of each normalized sample library to be pooled in a 1.5-mL microcentrifuge tube (see Note 3). Each library in the pooled mixture must have a different index or barcode.
8. Send an aliquot of the pooled libraries to be sequenced (see Note 6) or store at -20ºC. It is recommended that the libraries are shipped on dry ice.

3.11 Library Preparation for sRNA sequencing (LIB<500 and LIB<200 libraries)

The LIB<500 and LIB<200 libraries are prepared with the Illumina TruSeq™ Small RNA Sample Preparation Kit (Part Number 15004197 Rev. D May 2012 / http://support.illumina.com/documents/MyIllumina/b76a55b0-4ac7-4ee7-aa23-0e993cf7f467/TruSeq_SmallRNA_SamplePrep_Guide_15004197_D.pdf) (see Notes 3 and 5) with some modifications that are explained below. Read the *Best Practices* section from the manual carefully. If more than 8 samples are to be prepared, then we recommend using 250-µL PCR plates and caps for PCR strips that are compatible with the plates to perform the protocol.

1. Start the protocol at the section *Ligate 3' and 5' Adapters* with 100-400 ng of size-selected RNA, diluted in a maximum volume of 5 µL nuclease-free water (see Note 15).
2. Perform all steps from sections *Ligate 3' and 5' Adapters* and *Reverse Transcribe and Amplify.* At this point the amplified libraries are in a 50 µL volume.
3. Allow the AMPure XP Beads to come to room temperature prior to use. Vortex the beads until they are well dispersed, and then add 50 µL of the mixed AMPure XP Beads to each PCR amplified library. Gently pipette the entire volume up and down ten times to mix thoroughly.
4. Incubate the PCR tubes or plate at room temperature for 15 min.
5. Place the PCR tubes or plate on the magnetic stand at room temperature for 5 min or until the liquid appears clear.
6. Remove and discard 95 µL of the supernatant from each PCR tube or plate well.
7. With the PCR tubes or plate remaining on the magnetic stand, add 200 µL of freshly prepared 80% EtOH to each tube or well without disturbing the beads.

8. Incubate at room temperature for 30 s, then remove and discard all of the supernatant from each tube or well.
9. Repeat the last two steps once for a total of two 80% EtOH washes.
10. While keeping the PCR tubes or plate on the magnetic stand, let the samples air dry at room temperature for 15 min and then remove the tubes or plate from the magnetic stand.
11. Resuspend the dried pellet with 22.5 µL TE buffer. Gently pipette the entire volume up and down ten times to mix thoroughly.
12. Incubate at room temperature for 2 min.
13. Place the PCR tubes or plate on the magnetic stand at room temperature for 5 min or until the liquid appears clear.
14. Transfer 20 µL of the clear supernatant to a clean PCR tube or plate well.
15. Quantify the libraries using the Qubit® dsDNA HS BR Assay Kit on the Qubit® Fluorometer (Invitrogen) (see Note 17).
16. Check the size and purity of the sample by running a High Sensitivity DNA chip on the Bioanalyzer. The final product should be a broad band with an average size of approximately of 200 bp (see Figure 3B).
17. Adjust the concentration of the libraries to 10 nM using a 10 mM Tris-Cl (pH 8.5) solution containing 0.1% Tween 20.
18. Combine 10 µL of each normalized sample library to be pooled in a 1.5-mL microcentrifuge tube (see Note 3). Each library in the pooled mixture must have a different index or barcode.
19. Send an aliquot of the pooled libraries to be sequenced (see Notes 5 and 6) or store at -20ºC. It is recommended that the libraries are shipped on dry ice.

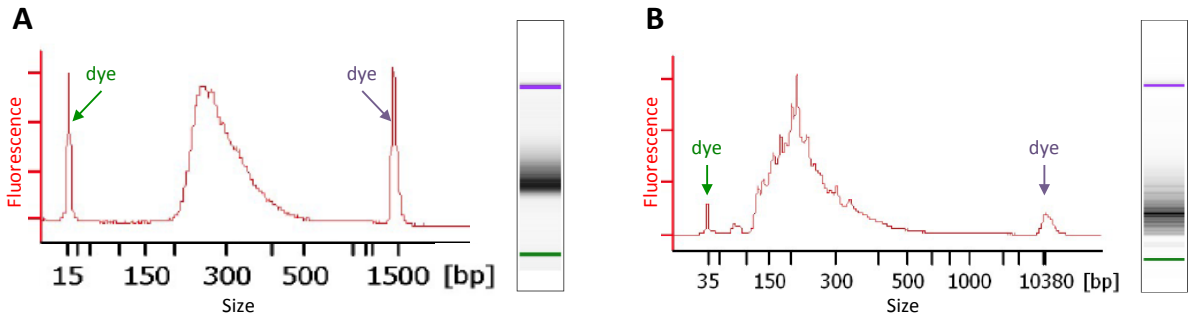


Figure 3. Typical size profiles of amplified libraries. A. The LIB>100 library analyzed with a DNA 1000 chip on the Bioanalyzer (Agilent). **B.** The LIB<500 library analyzed with a High Sensitivity DNA chip on the Bioanalyzer (Agilent). Libraries prepared using the LIB<200 protocol will have a very similar profile.

3.12 Read mapping

The first step of the data analysis pipeline is to map the reads against the reference genome. Note that cDNA libraries sequenced using a paired-end protocol will yield two FASTQ files per library, one containing the forward reads, and the other the reverse reads. In the LIB<500 and LIB<200 libraries, information corresponding to short transcripts (<100 nucleotides) is retained in the library construction. If the native read lengths are longer than the shortest cDNA fragments, then the read sequence will run through the cDNA fragment and extend into the adapter sequence. The 3'-end adapter sequences should therefore be trimmed away before mapping. There are several tools to trim reads, including the Flexbar tool that allows for flexible barcode detection and adapter removal. We map the adapter-trimmed reads onto the reference genome using Bowtie 2, which is a fast and memory-efficient tool for aligning short sequence reads onto reference genomes. The output of Bowtie 2 is the alignment of all the reads onto the reference genome in the generic Sequence Alignment/Map (SAM) format. The SAM format is a tab-delimited text format containing generic alignment information. For faster processing the SAM format can be converted into the equivalent binary format BAM.

3.13 Read coverage quantification

In order to calculate the transcription level across the genome at a single base resolution, the number of aligned reads for each position in the genome is counted. Thus, the global alignment stored in the SAM format needs to be converted into a list of the genomic positions together with the number of covering bases.

We use SAMtools for manipulating alignments in the SAM format. The SAM-formatted alignments are then converted into the BAM format, which is sorted to create a pileup file. Pileup format is a text-based format for summarizing the base calls of aligned reads at each covered position in the reference sequence. In the pileup file the information about the genomic location (chromosome and position) and the count of covering reads is stored in columns one, two and eight, respectively. Note that the pileup format only contains information about genomic positions that are covered by reads. It is therefore recommended to amend the pileup file to include uncovered positions designating the number of covering reads to be 0. If there is only one chromosome, the information about coverage depth along the genome can be stored in two columns, with the genomic position specified in the first column, and the number of covering reads given in the second column. This type of file will now be referred to as the coverage file.

3.14 Normalization and replicate merging

After determining the read coverage across the genome, the read counts in the coverage file are then normalized according to the total number of reads in the cDNA library. This is done in order to compare the level of transcription across different libraries. After normalization the normalized coverage files from replicate samples are merged by calculating the average read count at each position.

3.15 Full transcriptome analysis (library LIB>100)

If the library LIB>100 protocol has been followed, it is now possible to calculate the transcription levels of all already-annotated genes.

1. Using your normalized coverage files (not merged), calculate the average expression of each gene by dividing the normalized read coverage depth (the sum of coverage at each base of the gene) by the length of the transcript. It is necessary to have an annotation file containing all the annotated genes of your organism, their coordinates and length.
2. Analyze differential gene expression between the different samples, using ANOVA for example.

3.16 Identification of *trans*-encoded sRNAs (all libraries)

Most sRNAs characterized to date regulate gene expression by base pairing with mRNAs. These sRNAs fall into two broad classes: those encoded in *cis* on the DNA strand opposite the target RNA that share extended regions of perfect complementarity (*cis*-encoded antisense sRNAs), and those encoded in intergenic regions that share limited complementarity with their targets (*trans*-encoded sRNAs). In this section, we explain how to identify *trans*-encoded sRNAs. Although this type of analysis can be performed using all three types of libraries described here, a larger number of sRNAs will be detected in the LIB<500 and LIB<200 libraries because they are prepared from RNA that is size-selected and enriched for sRNAs (*6*). For the purpose of our analysis we define *trans*-encoded sRNAs as transcripts encoded entirely in the intergenic regions. Nonetheless, some *trans*-encoded sRNAs might extend into their neighboring genes. Also, some of the identified sRNAs might represent 5'-UTRs truncated before the start codon of their respective genes. In order to detect *trans*-encoded sRNAs in a high-throughput manner, we have employed an automated algorithm that detects transcripts starting and ending in intergenic regions.

The algorithm detects the sudden increases in read coverage depths in the intergenic regions by the following criteria: (a) A transcript is detected if the coverage depth increases above a cut-off that is set to discriminate noise caused by sequenced mRNA. Our cut-off was set based on the average expression intensities of forty-four previously validated sRNAs in *P. aeruginosa* (*6*). (b) The average transcriptional level (calculated as the normalized read coverage depth divided by the length of the transcript) of the detected transcript must be twice the previously defined cut-off to be considered further as an sRNA candidate. The two cut-off values are necessary and used together to identify sRNAs in that the second cut-off (b) ensures detection of very abrupt increases in read coverage depths whereas the first cut-off (a) enables the length of the transcripts to be defined more accurately.

After detecting transcripts, their transcriptional levels can be compared between samples as described in the previous section. The detected *trans*-encoded sRNAs can be visualized using a global alignment viewer such as Tablet (*8*) that may be helpful in more accurately defining the lengths of the sRNAs. Tablet visualizes the alignment using the sorted BAM format described in section 3.13. Tablet also requires an indexed BAM formatted alignment to be indexed. Indexing enables a more rapid alignment retrieval, and can be performed with SAMtools.

3.17 Identification of *cis*-encoded sRNAs (LIB<500 and LIB<200 libraries)

The next objective is detection of *cis*-encoded sRNAs. As these sRNAs are transcribed from the DNA strand opposite the target RNA, they can only be directly identified in strand-specific libraries like LIB<500 or LIB<200. The *cis*-encoded sRNAs are detected using an algorithm similar to the one used for detection of *trans*-encoded sRNAs. As the *cis*-encoded sRNAs are transcribed opposite to their respective genes, the analysis shoud be constrained to reads being in the opposite direction to the genes onto which they align.
Information on the direction of the aligned reads is contained in the second column of the SAM format. A number of different flags are used to describe the orientation of the single read and its corresponding mate, and flags “99”, “147”, “83” and “163” designate the read and its mate to be properly aligned within a distance as defined by the insert size. Read pairs with flags 99 or 147 originate from transcripts encoded on the plus strand, while read pairs with flags 163 or 83 originate from transcripts encoded on the minus strand. This allows us to only extract reads that map to a known gene, but have the opposite direction of the gene. The reads that originate from a gene encoded on the plus strand will have the flags 99 or 147. Thus, the reads that map to this gene and have the flags 163 or 83 need to

be extracted. The procedure is likewise for genes encoded on the minus strand, where the reads with flags 99 or 147 need to be extracted.

The extracted reads are placed into SAM formatted files that are processed into new pileup format files which can then be normalized and merged as previously described in sections 3.13 and 3.14. The algorithm described in section 3.16 is used in order to detect transcripts with sudden increases in read coverage depths. The difference in this case is that only transcripts that overlap genes are evaluated. The ends of mRNA molecules that do not encode proteins are termed untranslated regions (UTRs). These regions can contain regulatory elements for controlling gene expression and extend into the neighboring genes. Be aware of the fact that some *cis*-encoded sRNAs detected in your libraries might be UTRs from neighboring genes. Calculations of transcriptional levels, comparisons between samples, and visualization of *cis*-encoded sRNAs can be performed as previously described in sections 3.13 to 3.16.

4. NOTES

1. RNA is highly susceptible to degradation by ribonucleases. These enzymes are present on skin, laboratory equipment, and in dust. They are very stable and difficult to inactivate. For these reasons, it is important to follow best laboratory practices while preparing and handling RNA samples. Wear gloves and use sterile technique at all times. Reserve a set of pipettes for RNA work. Use sterile RNase-free filter pipette tips to prevent cross-contamination. Use RNase-free disposable plasticware. All reagents and solutions should be sterile and RNase-free where possible. Use RNase decontamination solution (such as RNase-ZAP) to decontaminate work surfaces and equipment prior to starting this protocol.
2. Although we have used precast gels to size-select transcripts shorter than 500 nt, 1-mm thick, 10% polyacrylamide TBE-Urea minigels can also be prepared. For each gel, mix 1 mL 10X TBE with 3.33 mL acrylamide: bis-acrylamide (37.5:1), 2.14 mL water, 4.2 g urea, 330 mL 1.6% ammonium persulfate, and 5 mL tetramethylethylenediamine (TEMED). The glass plates, combs, and chambers used should be compatible with the Mini-PROTEAN or other suitable electrophoresis system.
3. It is highly recommended to work with at least two biological replicates of every sample to be sequenced. This will enable a better comparison of the transcript expression levels between different samples. Thus, if there are twelve samples or conditions, and there are two replicates per sample, then there are twenty-four libraries to be sequenced. If multiplexed sequencing (sequencing different samples in the same lane)

is to be performed, then each library needs to be prepared using a different index or barcode. This will allow the different libraries to be pooled together and sequenced in one lane. Therefore, the number of indexes or barcodes needed is the same as the total number of libraries to be sequenced in the same lane. We have successfully sequenced up to twenty-four different RNA libraries in the same lane.

4. There are two different versions of the TruSeq™ RNA Sample Prep Kit v2 (Illumina), where each contains twelve different indexes or barcodes (twenty-four unique indexes or barcodes in total).
5. There are four different versions of the TruSeq™ Small RNA Sample Preparation Kit (Illumina), where each contains twelve different indexes or barcodes (forty-eight unique indexes or barcodes in total). This kit is designed for sequencing of eukaryotic microRNAs (miRNAs), which are around 22 nt in length. The libraries obtained following the kit manual will contain DNA fragments of around 150 bp, including the adapters. We describe modifications that allow using the TruSeq™ Small RNA Sample Preparation Kit for preparing libraries that enable sequencing of bacterial small RNAs, which normally range between 70 and 500 nt in length. The libraries prepared following our protocol will contain fragments that range from 100 to 400 bp, including the adapters. Make sure to inform your sequencing provider about these modifications.
6. The libraries described here were sequenced using the Illumina HiSeq2000 platform with a paired-end protocol and read lengths of 100 nt. On average, 16.7 million sequence reads were generated from each library, and of these 12.3 million were of sufficient quality to be mapped onto our reference genome (*6*). However, as sequencing technologies are rapidly evolving, sequencing using the latest available technology as well as the longest possible read length is recommended.
7. The RNA used to prepare the LIB<500 and LIB<200 libraries was treated with Tobacco Acid Pyrophosphatase (TAP) to ensure the inclusion of primary transcripts containing a 5´-triphosphate. Although primary and processed transcripts are not distinguishable in these libraries, it is possible to modify the approach with additional libraries to obtain this information. One approach is to discriminate primary from processed 5' ends by preparing two libraries from the same RNA sample. One library is prepared with untreated RNA and the other library is enriched for primary transcripts by treatment with Terminator 5´-Phosphate-Dependent Exonuclease (Epicentre), which degrades RNA with 5´-monophosphate but not 5´-triphosphate ends (*9*).
8. There are a number of protocols and kits for RNA isolation, but many of them do not quantitatively retain the smaller RNAs that are of interest here. This is the case for most of the kits that use spin columns. We use Trizol (Invitrogen) to isolate RNA because it is a quick method that allows retention of all RNA species. The protocol

described here can also be used to extract RNA from mucoid cells. If another method is used to extract RNA, it is important to ensure that small RNA species are retained by checking the sample with a Bioanalyzer. The RNA profile obtained after extraction with Trizol is shown in Figure 2A, and that obtained with a spin column kit (in this example the NucleoSpin RNA II kit from Macherey Nagel) is shown in Figure 2B.

9. One milliliter of Trizol is sufficient to extract RNA from 10 mL of *P. aeruginosa* culture harvested at OD_{600} of 0.5 to 3, where the typical yields of total RNA obtained are 40 to 150 µg. For other growth/harvest conditions or for bacteria other than *P. aeruginosa*, optimization of the RNA extraction method may be necessary.
10. The RNA can be stored as pellets in 75% ethanol at -20ºC or -80ºC for at least one year, or at 4ºC for at least one week. RNA storage in ethanol is highly recommended for long-term storage (weeks or months).
11. It is important that both the A_{260}/A_{280} and the A_{260}/A_{230} ratios of RNA solutions are close to 2.0. RNA has its absorbance maximum at 260 nm and this absorbance is not dependent on the pH of the solution. However, the absorbance of some contaminants (like proteins) is pH-dependent. This means that although the A_{260} reading of the RNA solution will remain the same at different pH values, the A_{280} reading will differ in a pH-dependent manner. Adjusting the pH of the RNA solution from approximately 5.5 to a slightly alkaline pH of 7.5-8.5 significantly increases RNA A_{260}/A_{280} ratios from approximately 1.5 to 2.0 (*10*). This should be taken into account when measuring the absorbance ratio of RNA dissolved in nuclease-free water. When isolating low amounts of RNA the A_{260}/A_{230} ratio tends to be lower. This indicates contamination of the RNA solution with chaotropic salts, phenol or protein.
12. The Bioanalyzer (Agilent) is a microfluidics-based platform for sizing, quantification and quality control of DNA, RNA, proteins and cells. Here it is used for checking the quality of RNA, and the size of DNA libraries. We recommend quantifying RNA and DNA using other technologies such as a low-volume spectrophotometer or Qubit® Fluorometer (Invitrogen). The Bioanalyzer allows for rapid analysis of RNA and DNA samples (30-40 minutes) with minimal sample consumption (1µL). The higher sensitivity afforded by the Bioanalyzer is the main advantage over gel analysis, where larger sample amounts are required.
13. Denaturing polyacrylamide-urea gels can also be used to size-select RNAs shorter than 200 nt (see Note 2). However, we find it easier and faster to use the *mir*Vana™ miRNA Isolation Kit for this purpose.
14. The MICROB*Express*™ Bacterial mRNA Enrichment Kit (Ambion) has been used because it employs subtractive hybridization to specifically deplete rRNAs from total RNA. In our libraries, the average percentage of read coverage corresponding to 23S,

16S and 5S rRNAs was 7.1%, 3.8%, and 0.15%, respectively (*6*). For bacteria other than *P. aeruginosa* and *E. coli* the oligos used to deplete rRNAs may need to be re-designed. There are also other commercially available kits to deplete rRNA. One of them uses exonuclease to preferentially degrade processed RNAs with 5′ monophosphate ends (mRNA-ONLY Prokaryotic mRNA Isolation kit, from Epicentre). Exonuclease degrades all processed RNAs and, although the majority of the processed RNAs are believed to be rRNAs and tRNAs, there may be other transcripts including sRNAs that are also processed. He *et al* (2010) compared the performance of the MICROB*Express*™ and mRNA-ONLY kits and found that only the former based on subtractive hybridization adequately preserved relative transcript abundance for quantitative analyses, whereas the latter based on exonuclease treatment greatly compromised mRNA abundance fidelity (*11*). Another commercially available kit based on subtractive hybridization contains capture oligonucleotides to remove 23S, 16S and 5S rRNAs (Ribo-Zero rRNA Removal Kit, Epicentre).

15. Concentration of RNA samples can be achieved via an ethanol precipitation. Follow step 10 from section *3.3 Treatment with DNase I* and resuspend in the appropriate volume of TE buffer or nuclease-free water.
16. Illumina recommends using Microseal 'B' adhesive seals (BioRad) to seal the PCR plates. However, we find that sealing the PCR plates using caps for compatible PCR strips is easier and safer.
17. In our experience the Qubit® Flurometer (Invitrogen) is a fast and easy alternative for library quantification relative to qPCR, which is recommended by Illumina.

ACKNOWLEDGEMENTS

This work was supported by the Novo Nordisk Foundation, Lundbeck Foundation and the Danish National Research Foundation.

REFERENCES

1. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009 Feb 20;136(4):615–28.
2. Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2011 Dec;3(12):a00379.
3. Vogel J, Sharma CM. How to find small non-coding RNAs in bacteria. Biol. Chem. 2005 Dec;386(12):1219–38.
4. Sharma C, Vogel J. Experimental approaches for the discovery and characterization of regulatory small RNA. Current Opinion in Microbiology. 2009;12(5):536–46.
5. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. Molecular Cell. 2011 Sep 16;43(6):880–91.
6. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Genome-wide identification of

novel small RNAs in Pseudomonas aeruginosa. Environmental Microbiology. 2012 Aug;14(8):2006–16.
7. Karpinets TV, Greenwood DJ, Sams CE, Ammons JT. RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. BMC Biol. 2006;4:30.
8. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, et al. Tablet--next generation sequence assembly visualization. Bioinformatics. 2010 Feb 1;26(3):401–2.
9. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature. 2010 Mar 11;464(7286):250–5.
10. Wilfinger W, Mackey K, Chomczynski P. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. BioTechniques. Eaton; 1997;22(3):474–81.
11. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG, et al. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. Nat Meth. 2010 Oct;7(10):807–12.

Article 3

Antisense small RNAs respond to osmotic, oxidative and antibiotic stress in *Pseudomonas aeruginosa*

Gómez-Lozano, M., Marvig, R. L., Tulstrup, M. V. L., Tribelli, P., & Molin, S. *Manuscript in preparation,* **2013**.

Antisense small RNAs respond to osmotic, oxidative and antibiotic stress in *Pseudomonas aeruginosa*

María Gómez-Lozano[1,2], Rasmus L. Marvig[1], Monica V. L. Tulstrup[2], Paula Tribelli[3], and Søren Molin*[1,2]

[1] *Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.*
[2] *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.*
[3] *Dpto. de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. IQUIBICEN-CONICET, Buenos Aires, Argentina.*

*To whom correspondence should be addressed (e-mail: sm@bio.dtu.dk).

ABSTRACT

The use of RNA-Seq technologies is revealing that bacterial RNA molecules are more than mRNAs, rRNAs or tRNAs. During the last years bacterial transcriptomes have been shown to contain intergenic RNAs, antisense RNAs, and untranslated regions, all capable of performing diverse regulatory functions. In this study we identified 249 antisense RNAs (asRNAs) in the opportunistic pathogen *Pseudomonas aeruginosa*. A considerable number of asRNAs were transcribed opposite to genes encoding membrane proteins and genes involved in the transport of small molecules, cell wall, lipopolysaccharide (LPS), and capsule biosynthesis. A substantial number of asRNAs significantly changed their expression under osmotic, oxidative and antibiotic stress, suggesting that asRNAs play a regulatory role during these conditions. In addition we have validated the existence of two antisense RNAs, opposite to the *dnaA* and *ponA* genes, which encode the chromosomal replication initiator protein DnaA and the penicillin-binding protein 1A, respectively.

INTRODUCTION

During the last years an increasing number of bacterial transcriptomes have been examined with tiling arrays and RNA sequencing (RNA-Seq) technologies, revealing that a significant number of protein-coding genes are also transcribed from the reverse complementary strand in a wide range of bacterial species (1-8). In general, overlapping transcription results in the generation of an antisense RNA (asRNA) whose size usually varies between various tens to hundreds of nucleotides (nt). The regulatory role of asRNAs was first reported more than 30 years ago in *Escherichia coli,* when the plasmid- and transposon-encoded asRNAs RNAI and CopA were found to negatively regulate plasmid copy number (9,10). However, until very recently, bacterial asRNAs were infrequently detected and therefore considered to be rare. By

2007, only about thirty bacterial asRNAs had been identified (reviewed in (11)). The number of reported asRNAs has lately exploded due to the use of tiling array and especially RNA sequencing (RNA-Seq) to study microbial transcriptomes. While the number of asRNAs detected in bacterial transcriptomes is rapidly growing, characterization of the physiological function of individual asRNAs is growing much slower. AsRNAs are known to affect the expression of the target gene by different mechanisms (reviewed in (12)). These include: (i) transcription interference, in which transcription from one promoter is suppressed by a second promoter located in the opposite strand, (ii) transcription attenuation, in which base pairing of the asRNA to the target mRNA causes premature transcription termination, (iii) changes in the target RNA stability, where the asRNA either promotes or blocks degradation or cleavage of the mRNA by ribonucleases, (iv) asRNAs can directly block ribosome binding, and (v) might indirectly impact ribosome binding (either positively or negatively) by affecting the target RNA structure. In addition, regulating the expression of the opposite gene is not the only function of certain asRNAs. Some asRNAs encode small proteins (13), and some have the potential to act on multiple targets in *trans* (14-18).

In this work, we used RNA sequencing to identify asRNAs in the human pathogen *Pseudomonas aeruginosa*. *P. aeruginosa* can cause severe infections, particularly in immuno-compromised and cystic fibrosis (CF) patients. Recently two independent studies identified antisense transcription in *P. aeruginosa*. One of the studies investigated strains PAO1 and PA14 at early stationary phase and found 60 asRNAs (19). Interestingly, a number of these asRNAs were strain-specific or showed strain-specific expression (19). The other study identified 384 antisense transcriptional start sites, but not full length asRNAs, in *P. aeruginosa* PA14 grown at 28ºC and at 37ºC (20). However, it is evident that, due to their regulatory functions, the expression of asRNAs occurs mostly in a transient manner and is dependent on specific environmental cues. Therefore, we have aimed to identify asRNAs in *P. aeruginosa* PAO1 grown in a range of 13 different conditions to ensure a comprehensive characterization of the *P. aeruginosa* asRNA-ome. The conditions studied include exponential and early-stationary phase, osmotic, oxidative and antibiotic stress. Exponentially growing cultures of *P. aeruginosa* PAO1 were exposed to sodium chloride (NaCl) and hydrogen peroxide (H_2O_2) during osmotic and oxidative stress conditions, respectively. The antibiotics used were four different ß-lactams, tobramycin, azithromycin, ciprofloxacin, colistin and tetracycline.

Recently, we showed that library preparation for RNA-Seq plays a fundamental role when aiming at identifying novel transcripts. Using three RNA-Seq library protocols with different sRNA abundance, we identified over 500 novel intergenic sRNAs in *P. aeruginosa* PAO1 (21,22). Although the use of three different libraries increased the number of novel transcripts identified, there were significant differences in the subset of transcripts detected

in each library (21). Libraries that were prepared with a size-selected fraction of RNA were more sensitive in the detection of intergenic sRNAs. In this study we identified 249 asRNAs using strand-specific libraries that contain transcripts shorter than 500 nt.

MATERIALS AND METHODS

Primers and adapters

Supplementary Table 1 lists all primers and adapters used in this study.

Plasmid construction

The plasmids used in this study are listed in Table 1. Plasmids for L-arabinose-inducible expression of asRNAs were constructed by amplification of plasmid pBAD33 (23) with primers F1-SacI and R1 producing a PCR product of 1250 bp (F1-SacI introduces a SacI restriction site). For amplification of the asRNA insert, genomic DNA from strain PAO1 was amplified with primers F2 and R2-XhoI (F2 starts with the sRNA +1 site and R2-XhoI introduces an XhoI restriction site). The 5´ends of primers F2 and R1 are complementary to each other. PCR synthesis by overlap extension (PCR-SOE) was carried out by combining the 2 previously amplified products and running PCR for 10 cycles, after which primers F1-SacI and R2-XhoI are added and the reaction continues for another 30 cycles. The obtained product was digested with SacI and XhoI and cloned into the MCS of plasmid mini-CTX2 (24).

Bacterial strains

The bacterial strains used in this study are listed in Table 1. The *Pseudomonas aeruginosa* PAO1 strain was used as wild-type strain. The *E. coli* CC118*λpir* strain was used as cloning host (25). Strains pBAD-*asdnaA,* pBAD-*asponA* and pBAD-control were created by integration of plasmids pBAD-*asdnaA,* pBAD-*asponA* and pBAD-control, respectively, at the *attB* site of strain PAO1, followed by Flp-mediated excision of the plasmid backbone, as described in (24). The first nucleotide of the transcribed products is the +1 of the sRNAs. This is an important consideration when cloning noncoding RNAs under the control of heterologous promoters, since altering the natural RNA start site may have unintended consequences for its function. The *attB* region was sequenced using primers P_{ser}-up and P_{ser}-down to make sure that the integration was successful and the plasmid backbone excised.

Table 1. Strains and plasmids used in this study.

Strain or plasmid	Relevant markers/genotype	Source or reference
Strains		
P. aeruginosa PAO1	Wild type	Laboratory strain
P. aeruginosa pBAD-*asdnaA*	pBAD-*asdnaA* fragment integrated at the *attP* site	This study
P. aeruginosa pBAD-*asponA*	pBAD-*asponA* fragment integrated at the *attP* site	This study
P. aeruginosa pBAD-control	pBAD-control fragment integrated at the *attP* site	This study
E. coli CC118*λpir*	*araD139 Δ(ara leu)7697 ΔlacX74 phoA20 galE galK thi rpsE rpoB argE(Am) recA1 λpir* lysogen	Herrero *et al* (1990)
Plasmids		
mini-CTX2	TcR; self-proficient integration vector with *tet*, Ω-FRT-attP- MCS, *ori*, *int*, *oriT* and *lacI*	Hoang *et al* (2000)
pBAD33	CmR, pBAD expression plasmid with pBR322	Guzman *et al* (1995)
pBAD-*asdnaA*	TcR; AsdnaA expression plasmid, *asdnaA* is controlled by the plasmid-borne PBAD promoter	This study
pBAD-*asponA*	TcR; AsponA expression plasmid, *asponA* is controlled by the plasmid-borne PBAD promoter	This study
pBAD-control	TcR; pBAD control plasmid, expresses 322 nt nonsense RNA followed by a terminator	This study
pFLP2	AmpR, source of Flp recombinase	Hoang *et al* (1998)

Growth conditions

Growth in Luria–Bertani (LB) broth (250 rpm, 37°C) or on LB plates at 37°C was used throughout this study.
Stress exposure experiments. Overnight cultures of *Pseudomonas aeruginosa* PAO1 were diluted to a starting OD_{600} of 0.01 and grown to an OD_{600} of 0.5, at which an antibiotic, 0.5 M NaCl or 1mM hydrogen peroxide was added. The final concentrations of antibiotics are shown in Table 2. Concentrated stock solutions of H_2O_2 and all antibiotics were prepared fresh daily. Cells were harvested 1 hour after the addition of antibiotics, NaCl and H_2O_2. Early-stationary phase cells were harvested from cultures grown to an OD_{600} of 3. RNA was extracted and used to make RNA-Seq libraries LIB>100 and LIB<500. Experiments were performed in duplicates.
Pulse-expression experiments. Overnight cultures of strains pBAD-*asdnaA* and pBAD-*asponA* were diluted to a starting OD_{600} of 0.01 and grown to an OD_{600} of 0.5, at which sRNA expression was induced with 1% L-arabinose. Cells were harvested after 30 min of the

induction. RNA was extracted and used to make RNA-Seq libraries LIB>100, which contain the full transcriptome. Two sets of negative-control conditions were used for comparison. In one, the control strain pBAD-control was induced with arabinose; in the second, cells containing the pBAD-*asdnaA* and pBAD-*asponA* constructs were grown without arabinose induction. Experiments were performed in duplicates. The samples were analyzed by RNA-seq using a strand-specific library that contains transcripts longer than 100 nt. These libraries were prepared using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre) following the manufacturer´s instructions. Briefly, rRNA-depleted RNA was fragmented and reverse transcribed using random primers containing a 5'-tagging sequence. The 5'-tagged cDNA was then tagged at its 3' end by the terminal-tagging reaction to yield di-tagged, single-stranded cDNA. Following purification, the di-tagged cDNA was amplified by PCR, which completes the addition of the Illumina adaptor sequences, amplifies the library for subsequent cluster generation and adds a barcode. The amplified RNA-seq library was purified with the Agencourt AMPure XP beads (Beckman Coulter Genomics).

MIC value calculation

MIC values of *P. aeruginosa* PAO1 were assessed using both the broth microdilution procedure and E-test strips.
Broth Microdilution. LB medium was added to all wells of a 96-well microtiter plate loaded with serially diluted antibiotics. Each well was inoculated with *P. aeruginosa* PAO1 at a final concentration of 5 x 10^5 cfu/ml. The plates were incubated for 24 hours (250 rpm, 37°C). Following incubation, the optical density of all wells was measured and the lowest concentration showing complete inhibition of growth was recorded as the MIC for that antibiotic. The experiments were performed in triplicates. Concentrated stock solutions of all antibiotics were prepared fresh daily.
E-Test. E-test (bioMérieux AB BIODISK) strips were used according to the manufacturer's instructions. LB plates were inoculated equal amount of *P. aeruginosa* PAO1 cells. After drying, the E-test strips were placed on the top of the plates. The MIC values were read after 16 h of incubation at 37ºC. The experiments were performed in triplicates.

RNA isolation

Harvested cells were mixed immediately with 0.2 volumes of STOP solution (95% ethanol, 5% phenol) and pelleted by centrifugation. Total RNA was extracted with Trizol (Invitrogen). Removal of DNA was carried out by treatment with DNase I (Fermentas) in combination with the RNase inhibitor RiboLock (Fermentas). The integrity of total RNA and DNA contamination were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

Removal of 23S, 16S and 5S rRNAs

The 23S, 16S and 5S rRNAs were removed by subtractive hybridization using the MICROBExpress kit (Ambion) with modifications as previously described (21) (22). Capture oligonucleotides complementary to the rRNAs of *P.aeruginosa* PAO1 were used (Supplementary Table S1). For preparation of library LIB>100 (described below), 5 µM of each capture oligonucleotide was used, for a total capture oligonucleotide concentration of 25 µM. Compared with the standard protocol, 25% more capture oligonucleotides and magnetic beads were used. For preparation of libraries LIB<500 (described below), 5 µM of the 5S rRNA capture oligonucleotide and 50% less magnetic beads were used relative to the standard protocol. Removal of rRNAs was confirmed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

Library preparation and RNA sequencing

Two different sequencing libraries (LIB>100 and LIB<500) were constructed as previously described (21,22). Each library type was prepared with RNA isolated from cells grown in duplicate for each condition studied. After each step the samples were validated using an Agilent 2100 Bioanalyzer (Agilent Technologies), and the final concentration was measured using a Qubit 2.0 Fluorometer (Invitrogen). The libraries were sequenced using the Illumina HiSeq2000 platform with a paired-end protocol and read lengths of 100 nt.
LIB>100 library. This type of library contains transcripts longer than 100 nt and was prepared using the TruSeq RNA Sample Preparation kit (Illumina) as previously described (21,22). Briefly, the rRNA-depleted RNA is fragmented using divalent cations under elevated temperature. The cleaved RNA fragments are copied into cDNA using reverse transcriptase and random primers, followed by second-strand cDNA synthesis using DNA polymerase I and RNase H. After this step, transcripts shorter than 100 nt are not retained due to the purification performed with the Agencourt AMPure XP beads (Beckman Coulter Genomics). The cDNA fragments then go through an end repair process, 3'-addition of single 'A' bases and adapter ligation. This is followed by product purification and PCR amplification to generate the final cDNA library.
LIB<500 library. This type of library is strand-specific and contains transcripts shorter than 500 nt. RNA size selection was performed by running total RNA on 10% polyacrylamide gels containing 10 M urea. Gel slices corresponding to RNAs up to 500 nt were excised, followed by elution of RNA in 0.4 M NaCl and precipitation with ethanol. The 5S rRNA was

depleted as previously described, followed by treatment with Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37°C for 90 min. Tobacco Acid Pyrophosphatase (TAP) is used to convert 5′-triphosphate RNA into 5′-monophosphate RNA, which is important for correct adapter ligation. This was followed by treatment with RNase III (Ambion) for 10 min at 37°C to fragment the RNA. RNase III fragments RNA into smaller pieces containing a 5′-phosphoryl group and a 3′-hydroxyl group, which is important for specific adapter ligation in the next step. Sequential ligation of RNA 3′ and 5′ adapters was performed using the adapters and enzymes from the TruSeq Small RNA Sample Preparation kit (Illumina). Next, reverse transcription followed by PCR amplification was performed to form cDNA constructs based on the RNA fragments ligated with 3′ and 5′ adapters, selectively enriching fragments with adapter molecules on both ends. The reverse transcription and subsequent PCR amplification were performed using the enzymes and reagents from the TruSeq Small RNA Sample Preparation kit (Illumina). Agencourt AMPure XP beads (Beckman Coulter Genomics) were used for the post-PCR clean-up. In this type of library the fragments corresponding to transcripts shorter than 100 nt are retained due to ligated RNA adapters that increase the length of the fragments by 125 nt.

Data analysis

Our analysis pipeline is described in detail in Gómez-Lozano et al (2013) (22). Briefly, reads were mapped onto the *P. aeruginosa* PAO1 genome using the Bowtie 2 short read aligner (26). Read alignments from Bowtie 2 were handled using SAMtools (27). In order to obtain normalized expression intensities of the read coverage depth at each position in the genome, the number of reads in each replicate was normalized according to the total number of reads in the library, and expression intensities from replicate samples were averaged. A custom-made script was designed to detect novel transcripts antisense to annotated genes. Only transcripts of at least 70 nt were considered further. The resulting transcripts from automatic classification were re-evaluated by visual inspection. The asRNAs were categorized depending on their position with respect to the opposite gene as “start-span” (the asRNA spans over the start of the antisense gene), “internal” (the asRNA start and end within the antisense gene), or “end-span” (the asRNA spans over the end of the gene). A two-sample Student’s *t*-test was performed on the average expression of the transcripts to determine those with differential expression between the conditions tested (*P*-value <0.05 and fold change >2).

Rapid Amplification of cDNA Ends (RACE)

Schematic representations of the 5´ and 3´ RACE protocols can be seen in Supplementary Figures 1 and 2. Supplementary Table 1 lists primers and adapters used.

5´ RACE. 2 µg rRNA-depleted RNA was incubated with 10 U of Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37ºC for 1 h to convert RNA 5´ triphosphates in 5´ monophosphates. The same amount of control RNA was incubated under the same conditions in the absence of the enzyme. Reactions were stopped by phenol-chloroform extraction, followed by etanol-sodium acetate precipitation. Precipitated RNAs were redissolved in water, mixed with 500 pmol of 5´ RNA adapter, heat-denatured at 95ºC for 5 min, then quick-chilled on ice. A short RNA adapter was ligated was ligated with 50 U of T4 RNA ligase (Thermo Scientific) at 37ºC for 1 h. Reactions were stopped by phenol-chloroform extraction, followed by ethanol-sodium acetate precipitation. Precipitated RNAs were re-dissolved in 20 µl water. Then 10 µl of 5´adapter-ligated RNA was reverse-transcribed using 2 pmol of primer complementary to the sRNA (5´-GSP1) and the Thermoscript RT-PCR system (Invitrogen) according to the manufacturer’s instructions. Reverse transcription (RT) was performed in three subsequent 20 min steps at 55ºC, 60ºC, and 65ºC, followed by treatment with RNase H. Primers were removed using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The products of RT were amplified using 10 pmol of another primer complementary to the sRNA (5´-GSP2) and 10 pmol the 5´ adapter-specific primer, together with the Maxima Hot Start PCR Master Mix (Thermo Scientific) according to the manufacturer’s instructions. Negative controls were performed using the 5´ adapter-ligated RNA as template. The PCR products were resolved and purified using E-Gel SizeSelect 2% Agarose gels (Invitrogen). Products were sequenced with 5´-GSP2 and 5´ adapter-specific primers by LGC Genomics GmbH (Germany).

3´ RACE. 2 µg rRNA-depleted RNA was dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) at 37ºC for 1 h. Reactions were stopped by phenol-chloroform extraction, followed by etanol-sodium acetate precipitation. Ligation was done as above with a 3´ RNA adapter with a 3´-inverted deoxythymidine (3´-idT). RT was carried out as described, but with 10 pmol of a single primer complementary to the 3´ RNA adapter. PCR amplification, band purification and sequence analysis was done as described above. All enzymatic treatments of RNA were performed in the presence of 20 units of RiboLock RNase Inhibitor (Thermo Scientific).

RESULTS AND DISCUSSION

Identification of antisense sRNAs.

Recently, we identified over 500 novel intergenic sRNAs (21) using three different RNA-Seq library preparations (22). Although the use of three different libraries increased the number of novel transcripts identified, there were significant differences in the subset of transcripts detected in each library, underscoring the importance of library preparation strategy and relative sRNA abundance for successful sRNA detection (21). In this study we identified 249 asRNAs using the library type LIB<500 under different stress conditions, including stationary phase, osmotic, oxidative and antibiotic stress. Libraries prepared using the LIB<500 protocol are strand-specific and contain transcripts shorter than 500 nt (21,22). Exponentially growing cultures of *P. aeruginosa* PAO1 were exposed to sodium chloride (NaCl) and hydrogen peroxide (H_2O_2) during osmotic and oxidative stress conditions, respectively. The antibiotics used were four different ß-lactams (piperacillin, ceftazidime, aztreonam, meropenem), the aminoglycoside tobramycin, the macrolide azithromycin, the fluoroquinolone ciprofloxacin, the polymyxin antibiotic colistin, and the polyketide tetracycline. All antibiotics used in this study, with the exception of tetracycline, are routinely used against *P. aeruginosa* infections (28-32). The concentrations of antibiotics, NaCl and H_2O_2 used in this study are shown in Table 2. The conditions were chosen due to their clinical relevance during *P. aeruginosa* infection in CF patients, in which the bacteria are exposed to high salt concentrations, reactive oxygen species (ROS) secreted by the host immune response, and aggressive antibiotic treatments. The CF lung is an osmotically stressful environment, due to the viscous, dehydrated mucus, cellular debris and electrolytes competing for a limited supply of water (33,34). A variety of studies have shown that *P. aeruginosa* encounters ROS in the lungs of CF patients due to the exaggerated, sustained and extended inflammatory response, characterized by influx of neutrophils and high concentrations of interleukin-8 (35-37). In addition, *P. aeruginosa*-infected CF patients are routinely treated with several types of antibiotics, because early and aggressive antipseudomonal treatment regimens are correlated with improved lung function and survival of CF patients (38).

Table 2. Concentration of antibiotics, NaCl and H_2O_2 used in the stress exposure experiments.

Condition	Abbreviation	MIC value (µg/ml)	Concentration added at OD=0.5
Control	-	-	-
Aztreonam	AZT	4	3xMIC
Ceftazidime	CEF	2	3xMIC
Ciprofloxacin	CIP	0.125	3xMIC
Meropenem	MER	1	3xMIC
Tetracyclin	TET	8	3xMIC
Tobramycin	TOB	1	3xMIC
Azithromycin	AZI	1.5	10xMIC
Colistin	CO	4	3xMIC
Piperacillin	PP	4	3xMIC
Hydrogen peroxyde	H_2O_2	-	1 mM
Sodium chloride	NaCl	-	0.5 M

The analysis of the data generated revealed the existence of 249 novel asRNAs. Supplementary Table 2 lists the coordinates of the asRNAs identified in this study, as well as the annotation of the opposite genes. Antisense transcripts can overlap with the 5' end, the 3' end, an internal part of the opposite gene, the whole gene, or even various contiguous genes that have the same direction. The asRNAs were categorized depending on their position with respect to the opposite gene as "start-span" (the asRNA spans over the start of the antisense gene), "internal" (the asRNA start and end within the antisense gene), or "end-span" (the asRNA spans over the end of the gene). Only asRNAs longer than 50 nt were taken into consideration. The longest detected asRNA is 581-nt long. Eighteen asRNAs overlapped with two contiguous genes transcribed from the opposite strand. Alternatively, antisense transcription can also be due to the overlap between 5' and 3' untranslated regions of genes encoded in the opposite direction. UTRs can contain regulatory elements for controlling gene expression and extend into the neighboring genes. Some asRNAs detected in RNA-Seq libraries might actually be UTRs from neighboring genes. To prevent the mis-annotation of UTRs as asRNAs, in this study we have not considered as asRNAs antisense transcription that finished close (less than 100 bp) to the start of a flanking gene having the same direction as the asRNA. We cannot rule out though that some of the novel asRNAs identified in this study might be novel coding genes encoded opposite to already-annotated coding regions. It is also possible that some of the asRNAs identified are the result of non-specific transcription and thus do not have physiological functions.

We next examined the distribution of antisense sRNAs and found that they appear to be homogenously distributed in the genome of *P. aeruginosa* PAO1 (Figure 1A). A recent study made by Wurtzel *et al* (2012) found that antisense sRNAs are more often found in some parts of the accessory genome of *P. aeruginosa* PA14 (20). Specifically, the authors identified 384 sites with overlapping transcription on the reverse strand, and found that the pathogenicity

island PAPI-1 is enriched in these kind of sites by 5-fold compared to the core genome (20). In this study we do not observe more concentration of antisense transcripts in the accessory than in the core genome. There are three reasons that might explain these differences between this study and the one from Wurtzel *et al* (2012). (i) The strain used in this study is *P. aeruginosa* PAO1, while Wurtzel *et al* (2012) used strain PA14 (20). Another study investigated the expression of sRNAs of strains PAO1 and PA14 at early stationary phase, and found 60 asRNAs (19). Interestingly, a number of these asRNAs were strain-specific or showed strain-specific expression (19). This also seems to be the case for intergenic sRNAs, whose sequences are sometimes not conserved between different *P. aeruginosa* strains (21). In addition, the PAPI-1 island is found in strain PA14 but not in PAO1, though it can be transferred by a type IV pilus (39). (ii) Wurtzel *et al* (2012) studied two conditions, growth at 28ºC and at 37ºC (20). In this study we sequenced samples from *P. aeruginosa* in 12 different conditions. Studying more conditions might lead to finding antisense transcription sites in more genomic locations. (iii) Due to their library preparation protocol, Wurtzel *et al* (2012) identified antisense transcription start sites, but the authors could not determined the length of the asRNAs (20). In this study, we identified 249 asRNAs with lengths ranging from 50 to 581 nt in length. We did not take into consideration neither asRNAs shorter than 50 nt, nor asRNAs that might be suspected to be UTRs from flanking genes. This might explain why in this study we identify considerably less asRNAs than in Wurtzel *et al* (2012).

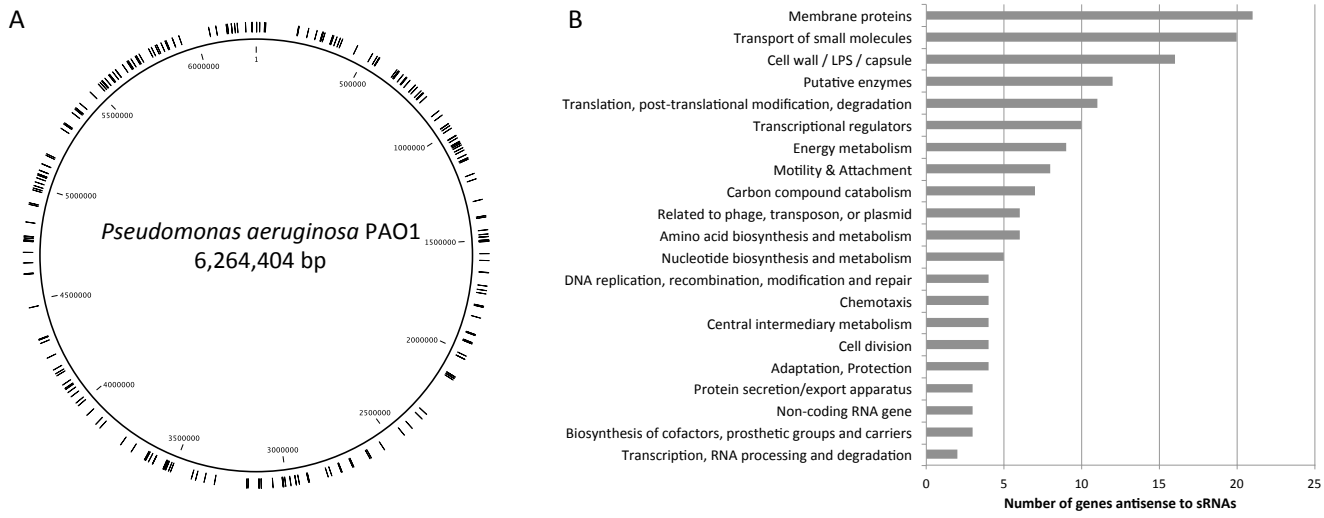


Figure 1. Distribution and targets of asRNAs. A. Distribution of asRNAs in the genome of *P. aeruginosa* PAO1. Each asRNA is marked by a black line. **B.** Categories of genes with asRNAs transcribed from the reverse strand.

Intergenic and antisense sRNAs respond to stress.

Figure 1B shows the classification of genes with asRNAs transcribed from the reverse strand in our experiments. A considerable number of asRNAs are transcribed opposite to genes encoding membrane proteins, genes involved in the transport of small molecules, cell wall, lipopolysaccharide (LPS), and capsule biosynthesis. This most likely reflects the conditions used in this study: stationary phase, oxidative, osmotic and antibiotic stress. In these conditions, fast and fine-tuned regulation of the membrane proteins and transport is essential. We hypothesize that the regulatory effects of asRNAs are important in probably all cellular processes, and that the categories of genes with antisense transcription regulation mainly depend on the conditions studied.

In addition to the 249 asRNAs identified, we recently identified over 500 intergenic sRNAs in *P. aeruginosa* PAO1 (21). Figure 2 represents the number of intergenic (left) and asRNAs (right) whose expression is significantly changed during the conditions tested. A substantial number of both intergenic and asRNAs significantly change their expression, suggesting that their regulatory effects are important during antibiotic, H_2O_2, NaCl treatment, and stationary phase. About half of the asRNAs detected in this study were differentially expressed during stationary phase. Supplementary Table 3 lists the asRNAs differentially transcribed in our conditions, as well as their fold-change in expression. We anticipate that these data will be important to understand the regulation of genes that show antisense transcription on the reverse strand. However, the level of transcription alone cannot explain the mechanism of action of asRNAs, and more experiments will be needed in order to determine how asRNAs regulate the opposite genes.

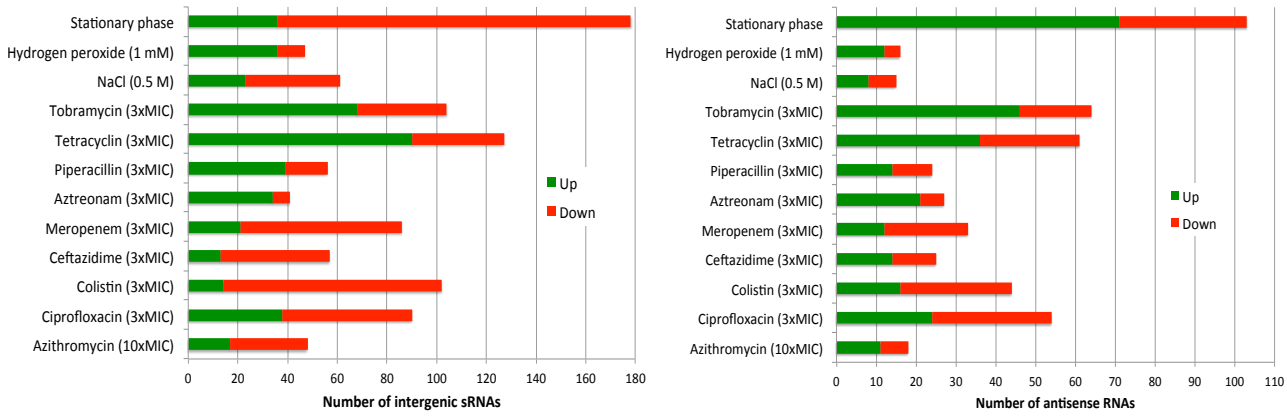


Figure 2. Differential expression of sRNAs. Number of intergenic (left) and asRNAs (right) whose expression is significantly changed during antibiotic, H_2O_2 or NaCl treatment.

Antisense sRNAs targeting *dnaA* and *ponA*

Two novel asRNAs were validated using 5´and 3´-RACE. The coordinates of both asRNAs were very similar between the RNA-Seq data and the RACE experiments, with differences of up to 19 nt between the two techniques. One asRNA, called from now on AsdnaA, is transcribed from the reverse strand of the *dnaA* gene, which encodes the chromosomal replication initiator protein DnaA (Figure 3). AsdnaA is highly expressed during stationary phase (8.7-fold) and upon ciprofloxacin (4.6-fold), aztreonam (2.9-fold) and H_2O_2 (2.1-fold) exposure. The other asRNA is called AsponA and its transcription is antisense to the *ponA* gene, encoding the penicillin-binding protein 1A (Figure 3). AsponA is upregulated upon aztreonam (2.9-fold), piperacillin (2.8-fold), and ciprofloxacin (2-fold) exposure, and downregulated during meropenem (2-fold), colistin (3-fold), tobramycin (5.4-fold), and tetracycline (5.6-fold) exposure.

To determine the effect of these two asRNAs on their opposite genes, pulse-expression experiments were carried out, in which the asRNAs are pulse-expressed with an inducible promoter and its immediate effects studied by RNA-Seq. Cells carrying the pBAD-*asdnaA* or the pBAD-*asponA* construct were grown to mid-log phase in LB and treated with 1% arabinose for 30 min. Two sets of negative-control conditions were used for comparison. In one, the control construct, pBAD-control was induced with arabinose; in the second, cells containing either the pBAD-*asdnaA* or pBAD-*asponA* construct were grown without arabinose induction. Total RNA was extracted from all cultures and processed for RNA-Seq analysis (see Materials and Methods). The gene expression profiles were compared in two ways: (i) induced pBAD-*asdnaA* or pBAD-*asponA*-containing cells were compared to induced pBAD-control-containing cells, and (ii) uninduced pBAD-*asdnaA* or pBAD-*asponA*-containing cells were compared to induced pBAD-*asdnaA* or pBAD-*asponA*-containing cells. The results for the two types of comparisons showed consistent results, with the exception of the expected differences for arabinose-inducible genes. It is worth mentioning that while pulse-expression experiments of asRNAs can provide important insights about their role, they can also cause artifactual effects that are not observed during endogenous expression of asRNAs.

AsdnaA base is encoded opposite to the second half of gene *dnaA* (Figure 3). Due to its location, we initially hypothesize that AsdnaA might (i) attenuate the transcription of *dnaA*, by prematurely terminating the transcription on the opposite strand, (ii) impact the stability of *dnaA* by either promoting or blocking cleavage by ribonucleases, or (iii) indirectly impact (either positively or negatively) ribosome binding by affecting the target mRNA structure. When pulse-expressing AsdnaA, the transcription of AsdnaA was increased by 130-fold after 30 min of induction. However the levels of the opposite gene *dnaA* were barely affected (1.4-fold more when pulse-expressing AsdnaA). This suggest that, if any, the physiological role of AsdnaA probably is to affect the level of translation of *dnaA*, but further investigations are needed to

prove it.

AsponA spans into the intergenic region overlapping with the predicted -10 and -35 σ-70 promoter elements (Figure 3). Due to the location of AsponA, we hypothesized that it probably belongs to the class of asRNAs that prevent the transcription or translation of the opposite gene, either by transcription interference or by directly blocking ribosome binding. Wurtzel *et al* (2012) reported the start of antisense transcription at the same coordinates of the *ponA* gene (20). The conservation across strains together with the condition-dependent expression of the AsponA gives weight to the conclusion that this asRNA has a physiological role. Unfortunately, we did not observe any increase in the levels of asRNA AsponA after arabinose induction, suggesting that in the conditions tested AsponA is not stable. There are several possible reasons for this. (i) We are not over-expressing the primary transcript of AsponA, as we could not determine it by 5´-RACE. In pulse-expression analysis, the first nucleotide of the transcribed product should be the +1 of the primary transcript to be pulse-expressed. This is an important consideration when cloning sRNAs under the control of heterologous promoters, since altering the natural RNA start site may have unintended consequences for its function. We were able to identify the primary transcript of AsdnaA by 5´-RACE. However we did not detect the primary transcript of AsponA, and this might also explain why we did not obtain stable AsponA transcription in the pulse-expression experiment. (ii) AsponA might only be stable in conditions in which is needed, such as aztreonam, piperacillin, ciprofloxacin and H_2O_2 exposure. In conditions where the expression of AsponA is deleterious it might rapidly get degraded. (iii) Some asRNAs interfere with the transcription of opposite gene. In this mechanism transcription from the antisense promoter blocks transcription from the gene promoter by preventing RNA polymerase from either binding or extending the mRNA encoded on the opposite strand. Transcription interference does not require base pairing between the asRNA and the mRNA, and it does not occur when the asRNA is provided in *trans*. If this is the case of AsponA, then ectopic expression of the asRNA in *trans* will not affect the expression of the opposite gene and it might also promote the rapid degradation of the asRNA. In conclusion, more experiments are required to conclude whether AsponA plays a physiological role in *P. aeruginosa*.

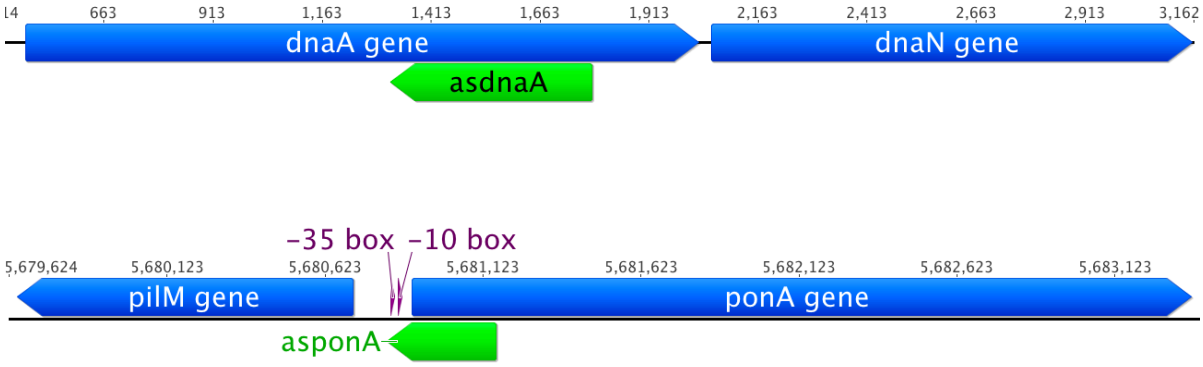


Figure 3. Antisense sRNAs AsdnaA and AsponaA. Antisense RNA AsdnaA (above) is encoded opposite of gene *dnaA*, which encodes the chromosomal replication initiator protein DnaA. Antisense RNA AsponA (below) is encoded opposite the beginning of gene *ponA*, encoding the penicillin-binding protein 1A, and spans into the intergenic region overlapping with the predicted -10 and -35 σ-70 promoter elements.

CONCLUDING REMARKS

In this work RNA-Seq was used to identify 249 asRNAs in the opportunistic pathogen *P. aeruginosa* under up to a total of 13 different conditions, including stationary phase, osmotic, oxidative and antibiotic stress. We provide the fold-change in expression of the asRNAs differentially transcribed in our conditions. A considerable number of asRNAs were transcribed opposite to genes encoding membrane proteins and genes involved in the transport of small molecules, cell wall, LPS, and capsule biosynthesis. A substantial number of asRNAs significantly changed their expression under the conditions tested, suggesting that asRNAs play a regulatory role during these conditions. Due to our library preparation protocol, only asRNAs ranging approximately from 50 to 500 nt were detected. Thus the number of asRNAs provided in this study is most likely an underestimate. Selected asRNAs could potentially be exploited as novel drug targets, and we anticipate that the data presented here will be important to understand the regulation of genes that show antisense transcription on the reverse strand. Riboswitches have already begun to be validated as anti-bacterial drug targets. For instance, a pyrimidine compound acting as a guanine riboswitch antagonist was shown to reduce *S. aureus* infection rates in mice (40). It might be possible to exploit similar strategies to target individual sRNAs by small compounds, or target mRNAs with synthetic antisense sRNAs, in order to combat bacterial infections in a species-specific manner.

Two asRNAs have been independently validated in this study, and their roles investigated by pulse-expression experiments. High expression levels of the asRNA asdnaA, which overlaps the *dnaA* gene, had no effect on the *dnaA* mRNA, and we hypothesize that either this asRNA has no physiological function or it affects the translation of the DnaA protein, by indirectly impacting (either positively or negatively) ribosome binding by affecting the target mRNA structure. The levels of the other asRNA, which

overlaps the *ponA* gene, were not increased after pulse-expression. We could not detect the primary transcript of AsponA, and this might be the reason for the failure of the pulse-expression. Alternatively this asRNA might not be stable under the conditions tested, or the transcription needs to be in *cis* to have an effect on the *ponA* mRNA. In conclusion, more experiments are required to rule out whether these two asRNAs play a physiological role in *P. aeruginosa*.

There are many questions that remain to be answered regarding the physiological roles of bacterial asRNAs. With the recent advances in transcriptomics, the hundreds of asRNAs now being reported in all bacteria studied need to be further validated and functionally characterized. We anticipate that understanding the molecular mechanisms of action of individual novel RNAs will be one of the biggest challenges of current RNA research.

ACKNOWLEDGEMENTS

This work was supported by the Novo Nordisk Foundation and the Danish National Research Agency.

REFERENCES

1. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature. 2010 Mar 11;464(7286):250–5.

2. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The Listeria transcriptional landscape from saprophytism to virulence. Nature. 2009;459(7249):950–6.

3. Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, Grills G, et al. Transcriptome Analysis of Pseudomonas syringae Identifies New Genes, Noncoding RNAs, and Antisense Activity. Journal of Bacteriology. 2010 May 1;192(9):2359–72.

4. Guell M, Van Noort V, Yus E, Chen W-H, Leigh-Bell J, Michalodimitrakis K, et al. Transcriptome Complexity in a Genome-Reduced Bacterium. Science. 2009 Nov 27;326(5957):1268–71.

5. Dornenburg JE, DeVita AM, Palumbo MJ, Wade JT. Widespread antisense transcription in Escherichia coli. mBio. 2010 Apr;1(1).

6. Albrecht M, Sharma CM, Reinhardt R, Vogel J, Rudel T. Deep sequencing-based discovery of the Chlamydia trachomatis transcriptome. Nucleic Acids Research. 2010 Jan;38(3):868–77.

7. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, et al. Genome-wide antisense transcription drives mRNA processing in bacteria. Proc Natl Acad Sci USA. 2011 Dec 13;108(50):20172–7.

8. Rasmussen S, Nielsen HB, Jarmer H. The transcriptionally active regions in the genome of Bacillus subtilis. Molecular Microbiology. Wiley Online Library; 2009;73(6):1043–57.

9. Stougaard P, Molin S, Nordström K. RNAs involved in copy-number control and incompatibility of plasmid R1. Proceedings of the National Academy of Sciences. 1981 Oct;78(10):6008–12.

10. Tomizawa J, Itoh T, Selzer G, Som T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proceedings of the National Academy of Sciences. 1981 Mar;78(3):1421–5.

11. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Current Opinion in Microbiology. 2007 Apr;10(2):102–9.

12. Thomason MK, Storz G. Bacterial Antisense RNAs: How Many Are There, and What Are They Doing? *. Annu. Rev. Genet. 2010 Dec;44(1):167–88.

13. Silby MW, Levy SB. Overlapping protein-encoding genes in Pseudomonas fluorescens Pf0-1. PLoS Genet. 2008 Jun;4(6):e1000094.

14. Sayed N, Jousselin A, Felden B. A cis-antisense RNA acts in trans in Staphylococcus aureus to control translation of a human cytolytic peptide. Nat. Struct. Mol. Biol. 2012 Jan;19(1):105–12.

15. Arnvig KB, Young DB. Identification of small RNAs in Mycobacterium tuberculosis. Molecular Microbiology. 2009 Aug;73(3):397–408.

16. Mandin P, Gottesman S. Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. EMBO J. 2010 Sep 15;29(18):3094–107.

17. Papenfort K, Said N, Welsink T, Lucchini S, Hinton JCD, Vogel J. Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. Molecular Microbiology. 2009 Oct;74(1):139–58.

18. Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P. Identification of new noncoding RNAs in Listeria monocytogenes and prediction of mRNA targets. Nucleic Acids Research. 2007;35(3):962–74.

19. Ferrara S, Brugnoli M, De Bonis A, Righetti F, Delvillani F, Dehò G, et al. Comparative Profiling of Pseudomonas aeruginosa Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs. PLoS ONE. 2012;7(5):e36553.

20. Wurtzel O, Yoder-Himes DR, Han K, Dandekar AA, Edelheit S, Greenberg EP, et al. The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. PLoS Pathog. 2012 Sep;8(9):e1002945.

21. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Genome-wide identification of novel small RNAs in Pseudomonas aeruginosa. Environmental Microbiology. 2012 Aug;14(8):2006–16.

22. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Identification of bacterial small RNAs by RNA sequencing. Methods in Pseudomonas aeruginosa. In press; 2013.

23. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology. 1995 Jul;177(14):4121–30.

24. Hoang TT, Kutchma AJ, Becher A, Schweizer HP. Integration-Proficient Plasmids for Pseudomonas aeruginosa: Site-Specific Integration and Use for Engineering of Reporter and Expression Strains. Plasmid. 2000 Jan;43(1):59–72.

25. Herrero M, de Lorenzo V, Timmis KN. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. Journal of Bacteriology. Am Soc Microbiol; 1990;172(11):6557–67.

26. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012 Apr;9(4):357–9.

27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.

28. Saiman L, Marshall BC, Mayer-Hamblett N, Burns JL, Quittner AL, Cibene DA, et al. Azithromycin in patients with cystic fibrosis chronically infected with Pseudomonas aeruginosa: a randomized controlled trial. JAMA. 2003 Oct 1;290(13):1749–56.

29. Doring G, Conway SP, Heijerman HG, Hodson ME, Hoiby N, Smyth A, et al. Antibiotic therapy against Pseudomonas aeruginosa in cystic fibrosis: a European consensus. Eur. Respir. J. 2000. pp. 749–67.

30. Aronoff SC, Klinger JD. In vitro activities of aztreonam, piperacillin, and ticarcillin combined with amikacin against amikacin-resistant Pseudomonas aeruginosa and P. cepacia isolates from children with cystic fibrosis. Antimicrobial agents and chemotherapy. Am Soc Microbiol; 1984;25(2):279–80.

31. Szaff M, Hoiby N, Flensborg EW. Frequent antibiotic therapy improves survival of cystic fibrosis patients with chronic Pseudomonas aeruginosa infection. Acta Paediatr Scand. 1983 Sep;72(5):651–7.

32. Høiby N. Recent advances in the treatment of Pseudomonas aeruginosa infections in cystic fibrosis. BMC Medicine. BioMed Central Ltd; 2011 Apr 4;9(1):32.

33. Chernick WS, Barbero GJ. Composition of tracheobronchial secretions in cystic fibrosis of the pancreas and bronchiectasis. Pediatrics. 1959 Nov;24:739–45.

34. Matthews LW, Spector S, Lemm J. Studies on pulmonary secretions. The over-all chemical composition of pulmonary secretions from patients with cystic fibrosis, bronchiectasis, and laryngectomy. The American review of Respiratory disease. 1963 Aug;88:199–204.

35. Hull J, Vervaart P, Grimwood K, Phelan P. Pulmonary oxidative stress response in young children with cystic fibrosis. Thorax. 1997 Jun;52(6):557–60.

36. Brown RK, Kelly FJ. Evidence for increased oxidative damage in patients with cystic fibrosis. Pediatr. Res. 1994 Oct;36(4):487–93.

37. Cantin A. Cystic fibrosis lung inflammation: early, sustained, and severe. American Journal of Respiratory and Critical Care Medicine. 1995 Apr;151(4):939–41.

38. Frederiksen B, Lanng S, Koch C, Hoiby N. Improved survival in the Danish center-treated cystic fibrosis patients: results of aggressive treatment. Pediatr. Pulmonol. 1996 Mar;21(3):153–8.

39. Carter MQ, Chen J, Lory S. The Pseudomonas aeruginosa pathogenicity island PAPI-1 is transferred via a novel type IV pilus. Journal of Bacteriology. 2010 Jul;192(13):3249–58.

40. Mulhbacher J, Brouillette E, Allard M, Fortier L-C, Malouin F, Lafontaine DA. Novel riboswitch ligand analogs as selective inhibitors of Guanine-related metabolic pathways. PLoS Pathog. 2009 Dec 31;6(4):e1000865–5.

Antisense small RNAs respond to osmotic, oxidative and antibiotic stress in *Pseudomonas aeruginosa*

María Gómez-Lozano[1,2], Rasmus L. Marvig[1], Monica V. L. Tulstrup[2], Paula Tribelli[3], and Søren Molin*[1,2]

[1] *Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.*
[2] *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.*
[3] *Dpto. de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. IQUIBICEN-CONICET, Buenos Aires, Argentina.*

Supplementary material

Supplementary Table 1. Primers and adapters used in the study.

Primer	Sequence (5´->3´)	Use
23S-1954	AAAAAAAAAAAAAAAAAAACTTACCCGACAAGGAATTTCGC	Removal of 23S rRNA (provided in the MICROBExpress kit, Ambion)
23S-2511	AAAAAAAAAAAAAAAAAAGAGCCGACATCGAGGTGCCAAAC	Removal of 23S rRNA (provided in the MICROBExpress kit, Ambion)
16S-807	AAAAAAAAAAAAAAAAAATGGACTACCAGGGTATCTAATCC	Removal of 16S rRNA (provided in the MICROBExpress kit, Ambion)
16S-1114	AAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTACGGGACTT	Removal of 16S rRNA (provided in the MICROBExpress kit, Ambion)
5S	AAAAAAAAAAAAAAAAAAGCGTTTCACTTCTGAGTTCGGCA	Removal of 5S rRNA
F1-SacI	GAGAGAGCTCGGCCGCTTCTAGAGTTATGA	Plasmids pBAD-*asdnaA,* pBAD-*asponA* and pBAD-control construction
R1_asdnaA	GCACCGTCGCCGAGTACTACATGGAGAAACAGTAGAGAGT	Plasmid pBAD-*asdnaA* construction
F2_asdnaA	GTAGTACTCGGCGACGGTGCGCTGGATGTTGTCGATGCTG	Plasmid pBAD-*asdnaA* construction
R2-XhoI_asdnaA	ACATCTCGAGGAAGCACACCAGCTATCTCA	Plasmid pBAD-*asdnaA* construction
R1_asponA	CCGCCGAGGACGACAATTTCATGGAGAAACAGTAGAGAGT	Plasmid pBAD-*asponA* construction
F2_asponA	GAAATTGTCGTCCTCGGCGGACAGCAGGGCATGGATGAAG	Plasmid pBAD-*asponA* construction
R2-XhoI_asponA	ACATCTCGAGTATGAGCCCTAGCACGACCA	Plasmid pBAD-*asponA* construction
R1_control	ATCAGGGCGTTCTTGATAAAATGGAGAAACAGTAGAGAGT	Plasmid pBAD-control construction
F2_control	TTTATCAAGAACGCCCTGATTCAGGACGGAGCGATAGATA	Plasmid pBAD-control construction
R2-XhoI_control	ACATCTCGAGTGGAATAGTGATGGCCACCA	Plasmid pBAD-control construction
Pser_up	CGAGTGGTTTAAGGCAACGGTCTTGA	Sequencing of pBAD-*asdnaA, pBAD-asponA* and pBAD-control
Pser_down	AGTTCGGCCTGGTGGAGCAACTCG	Sequencing of pBAD-*asdnaA, pBAD-asponA* and pBAD-control
5´RNA adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA	5´-RACE
5´RNA adapter-specific primer	GCTGATGGCGATGAATGAACACTG	5´-RACE
5´-GSP1_asdnaA	TGGAAGAGCGGCTGAAATCCC	5´-RACE *asdnaA*
5´-GSP2_asdnaA	GCTTCGGCTGGGGCCTGAC	5´-RACE *asdnaA*
5´-GSP1_asponA	AGATGCGCCGGACGCCGATC	5´-RACE *asponA*
5´-GSP2_asponA	CTTCGCCGACATCCCGCAGGA	5´-RACE *asponA*
3´RNA adapter E1	UUCACUGUUCUUAGCGGCCGCAUGCUC-idT	3´-RACE
3´RNA adapter E1 primer	CATGCGGCCGCTAAGAACAGTGA	3´-RACE
3´-GSP1_asdnaA	ATCAGGATCGCCACCCGGGTTT	3´-RACE *asdnaA*
3´-GSP2_asdnaA	TTCCGGCGGCTCGACGGCCACCGT	3´-RACE *asdnaA*
3´-GSP1_asponA	GAAATTGTCGTCCTCGGCGGACA	3´-RACE *asponA*
3´-GSP2_asponA	CAGGGCATGGATGAAGTCCTG	3´-RACE *asponA*

Supplementary Table 2. Annotation of antisense sRNAs.

Name	Start*	End*	Length*	Type of sRNA**	Antisense gene***	Direction gene	Gene name	Gene function	Gene pathway
asdnaA****	1320	1784	465	internal	PA0001	>	*dnaA*	chromosomal replication initiator protein DnaA	DNA replication, recombination, modification and repair
as3	15121	15215	95	end-span	PA0011	>		probable 2-OH-lauroyltransferase	Cell wall / LPS / capsule
as3	15121	15215	95	start-span	PA0012	>		hypothetical protein	Hypothetical, unclassified, unknown
as4	35866	35974	109	end-span	PA0033	<	*hptC*	Histidine phosphotransfer protein HptC	
as5	40775	40925	151	end-span	PA0039	>		hypothetical protein	Hypothetical, unclassified, unknown
as6	182516	182617	102	end-span	PA0158	>	*triC*	Resistance-Nodulation-Cell Division (RND) triclosan efflux transporter, TriC	Antibiotic resistance and susceptibility; Transport of small molecules
as7	184720	184861	142	internal	PA0162	>	*opdC*	histidine porin OpdC	Membrane proteins; Transport of small molecules
as8	226915	226976	62	internal	PA0196	>	*pntB*	pyridine nucleotide transhydrogenase, beta subunit	Transport of small molecules; Energy metabolism
as9	263722	263909	188	internal	PA0234	>		hypothetical protein	Hypothetical, unclassified, unknown
as10	288493	288639	147	internal	PA0257	<		hypothetical protein	Related to phage, transposon, or plasmid
as11	290529	290665	137	internal	PA0259	<		hypothetical protein	Hypothetical, unclassified, unknown
as12	291428	291543	116	internal	PA0260	<		hypothetical protein	Hypothetical, unclassified, unknown; Membrane proteins
as13	299094	299181	88	internal	PA0264	<		hypothetical protein	Hypothetical, unclassified, unknown
as14	334305	334355	51	internal	PA0296	<	*spuI*	Glutamylpolyamine synthetase	Putative enzymes; Carbon compound catabolism
as15	338651	338736	86	internal	PA0300	>	*spuD*	polyamine transport protein	Transport of small molecules
as16	353325	353458	134	internal	PA0313	<		L-cysteine transporter of ABC system YecS	Membrane proteins; Transport of small molecules
as17	368517	368669	153	internal	PA0328	<		hypothetical protein	Hypothetical, unclassified, unknown
as18	379215	379350	136	internal	PA0337	>	*ptsP*	phosphoenolpyruvate-protein phosphotransferase PtsP	Transport of small molecules
as19	448929	449284	356	internal	PA0407	<	*gshB*	glutathione synthetase	Amino acid biosynthesis and metabolism; Biosynthesis of cofactors, prosthetic groups and carriers
as20	521726	521838	113	internal	PA0461	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as21	544658	544778	121	internal	PA0483	>		probable acetyltransferase	Putative enzymes
as22	554963	555078	116	internal	PA0495	>		hypothetical protein	Hypothetical, unclassified, unknown
as23	557312	557569	258	end-span	PA0498	<		hypothetical protein	Hypothetical, unclassified, unknown
as24	646779	646892	114	internal	PA0588	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as25	654775	654912	138	internal	PA0595	<	*ostA*	organic solvent tolerance protein OstA precursor	Adaptation, Protection
as26	659114	659270	157	internal	PA0599	<		hypothetical protein	Hypothetical, unclassified, unknown
as27	680736	680853	118	internal	PA0620	>		probable bacteriophage protein	Related to phage, transposon, or plasmid
as28	697609	697703	95	internal	PA0641	>		probable bacteriophage protein	Related to phage, transposon, or plasmid
as29	701233	701489	257	end-span	PA0645	>		hypothetical protein	Related to phage, transposon, or plasmid
as29	701233	701489	257	start-span	PA0646	>		hypothetical protein	Related to phage, transposon, or plasmid
as30	718785	718934	150	internal	PA0667	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as31	748895	749047	153	internal	PA0689	>		hypothetical protein	Hypothetical, unclassified, unknown; Transport of small molecules
as32	749460	749562	103	internal	PA0689	>		hypothetical protein	Hypothetical, unclassified, unknown; Transport of small molecules
as33	761761	761823	63	internal	PA0690	>		hypothetical protein	Hypothetical, unclassified, unknown
as34	784213	784303	91	internal	PA0713	>		hypothetical protein	Hypothetical, unclassified, unknown
as35	795503	795648	146	internal	PA0727	>		hypothetical protein from bacteriophage Pf1	Hypothetical, unclassified, unknown; Related to phage, transposon, or plasmid
as36	798624	798854	231	start-span	PA0730	<		probable transferase	Putative enzymes
as37	817964	818095	132	start-span	PA0750	>	*ung*	uracil-DNA glycosylase	DNA replication, recombination, modification and repair
as38	898637	898918	282	end-span	PA0823	<		hypothetical protein	Hypothetical, unclassified, unknown
as38	898637	898918	282	start-span	PA0822	<		hypothetical protein	Hypothetical, unclassified, unknown
as39	921673	921840	168	end-span	PA0845	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as40	967079	967178	100	internal	PA0884	>		probable C4-dicarboxylate-binding periplasmic protein	Transport of small molecules
as41	967184	967283	100	internal	PA0884	>		probable C4-dicarboxylate-binding periplasmic protein	Transport of small molecules
as42	977508	977658	151	internal	PA0894	<		hypothetical protein	Hypothetical, unclassified, unknown
as43	990194	990309	116	internal	PA0904	>	*lysC*	aspartate kinase alpha and beta chain	Amino acid biosynthesis and metabolism
as44	1026460	1026563	104	internal	PA0936	<	*lpxO2*	lipopolysaccharide biosynthetic protein LpxO2	Putative enzymes; Cell wall / LPS / capsule
as45	1027896	1028083	188	end-span	PA0937	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as46	1039125	1039255	131	end-span	PA0952	<		hypothetical protein	Hypothetical, unclassified, unknown
as47	1056103	1056248	146	internal	PA0972	>	*tolB*	TolB protein	Transport of small molecules
as48	1060351	1060676	326	end-span	PA0977	<		hypothetical protein	Hypothetical, unclassified, unknown
as49	1066483	1066610	128	internal	PA0985	<	*pyoS5*	pyocin S5	Membrane proteins; Secreted Factors (toxins, enzymes, alginate)
as50	1073865	1074007	143	end-span	PA0992	>	*cupC1*	fimbrial subunit CupC1	Motility & Attachment
as50	1073865	1074007	143	start-span	PA0993	>	*cupC2*	chaperone CupC2	Motility & Attachment; Chaperones & heat shock proteins
as51	1087094	1087222	129	start-span	PA1003	<	*mvfR*	Transcriptional regulator MvfR	Biosynthesis of cofactors, prosthetic groups and carriers; Transcriptional regulators
as52	1098349	1098557	209	internal	PA1015	<		probable transcriptional regulator	Transcriptional regulators
as53	1119615	1119886	272	end-span	PA1032	<	*quiP*	QuiP	Carbon compound catabolism; Putative enzymes; Antibiotic resistance and susceptibility
as54	1135102	1135286	185	end-span	PA1047	>		probable esterase	Putative enzymes
as55	1135776	1135875	100	internal	PA1048	>		probable outer membrane protein precursor	Membrane proteins; Transport of small molecules
as56	1141250	1141364	115	start-span	PA1053	>		conserved hypothetical protein	Membrane proteins
as57	1207482	1207531	50	internal	PA1113	>		probable ATP-binding/permease fusion ABC transporter	Membrane proteins; Transport of small molecules
as58	1232028	1232169	142	start-span	PA1140	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as59	1323807	1323918	112	internal	PA1221	<		hypothetical protein	Hypothetical, unclassified, unknown
as60	1394817	1395002	186	internal	PA1284	<		probable acyl-CoA dehydrogenase	Putative enzymes
as61	1400645	1400720	76	end-span	PA1289	<		hypothetical protein	Hypothetical, unclassified, unknown
as62	1455941	1456046	106	internal	PA1342	<		probable binding protein component of ABC transporter	Transport of small molecules
as63	1474290	1474393	104	start-span	PA1362	>		hypothetical protein	Hypothetical, unclassified, unknown
as64	1484863	1485037	175	internal	PA1370	>		hypothetical protein	Hypothetical, unclassified, unknown
as65	1485541	1485595	55	internal	PA1370	>		hypothetical protein	Hypothetical, unclassified, unknown
as66	1499584	1499684	101	internal	PA1382	>		probable type II secretion system protein	Protein secretion/export apparatus
as67	1499690	1499742	53	internal	PA1382	>		probable type II secretion system protein	Protein secretion/export apparatus
as68	1502841	1502987	147	internal	PA1383	>		hypothetical protein	Hypothetical, unclassified, unknown
as69	1503570	1503696	127	internal	PA1384	>	*galE*	UDP-glucose 4-epimerase	Nucleotide biosynthesis and metabolism; Carbon compound catabolism; Central intermediary metabolism
as70	1503843	1503937	95	internal	PA1384	>	*galE*	UDP-glucose 4-epimerase	Nucleotide biosynthesis and metabolism; Carbon compound catabolism; Central intermediary metabolism
as71	1504553	1504676	124	end-span	PA1384	>	*galE*	UDP-glucose 4-epimerase	Nucleotide biosynthesis and metabolism; Carbon compound catabolism; Central intermediary metabolism
as71	1504553	1504676	124	start-span	PA1385	>		probable glycosyl transferase	Cell wall / LPS / capsule
as72	1504788	1504927	140	internal	PA1385	>		probable glycosyl transferase	Cell wall / LPS / capsule
as73	1506107	1506208	102	internal	PA1386	>		probable ATP-binding component of ABC transporter	Transport of small molecules
as74	1558500	1558752	253	internal	PA1430	>	*lasR*	transcriptional regulator LasR	Adaptation, Protection; Transcriptional regulators
as75	1588434	1588549	116	internal	PA1458	>		probable two-component sensor	Chemotaxis; Two-component regulatory systems
as76	1639661	1639805	145	end-span	PA1510	<		hypothetical protein	Hypothetical, unclassified, unknown
as76	1639661	1639805	145	start-span	PA1509	<		hypothetical protein	Hypothetical, unclassified, unknown
as77	1640773	1640948	176	internal	PA1510	<		hypothetical protein	Hypothetical, unclassified, unknown
as78	1698424	1698534	111	internal	PA1561	<	*aer*	aerotaxis receptor Aer	Adaptation, Protection; Chemotaxis
as79	1717815	1717930	116	end-span	PA1578	>		hypothetical protein	Hypothetical, unclassified, unknown
as80	1732127	1732265	139	end-span	PA1589	>	*sucD*	succinyl-CoA synthetase alpha chain	Energy metabolism
as81	1785747	1785881	135	internal	PA1640	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as82	1790530	1790650	121	internal	PA1645	<		hypothetical protein	Hypothetical, unclassified, unknown
as83	1838134	1838246	113	internal	PA1688	>		hypothetical protein	Hypothetical, unclassified, unknown
as84	1911985	1912118	134	internal	PA1767	<		hypothetical protein	Membrane proteins
as85	1912994	1913099	106	end-span	PA1769	<		conserved hypothetical protein	Hypothetical, unclassified, unknown

as86	1913728	1913919	192	start-span	PA1769	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as87	1916046	1916172	127	internal	PA1770	>	*ppsA*	phosphoenolpyruvate synthase	Energy metabolism; Carbon compound catabolism; Central intermediary metabolism
as88	1949036	1949142	107	internal	PA1797	<		hypothetical protein	Hypothetical, unclassified, unknown
as89	1953270	1953371	102	internal	PA1800	>	*tig*	trigger factor	Cell division; Chaperones & heat shock proteins
as90	1954335	1954415	81	internal	PA1801	>	*clpP*	ClpP	Chaperones & heat shock proteins; Cell wall / LPS / capsule
as91	1954420	1954469	50	internal	PA1801	>	*clpP*	ClpP	Chaperones & heat shock proteins; Cell wall / LPS / capsule
as92	2012282	2012454	173	internal	PA1852	<		hypothetical protein	Hypothetical, unclassified, unknown
as93	2027111	2027193	83	internal	PA1866	>		hypothetical protein	Hypothetical, unclassified, unknown
as94	2027457	2027571	115	internal	PA1866	>		hypothetical protein	Hypothetical, unclassified, unknown
as95	2113335	2113442	108	internal	PA1933	>		probable hydroxylase large subunit	Carbon compound catabolism
as96	2113618	2113734	117	internal	PA1933	>		probable hydroxylase large subunit	Carbon compound catabolism
as97	2113817	2113927	111	internal	PA1933	>		probable hydroxylase large subunit	Carbon compound catabolism
as98	2120607	2120921	315	internal	PA1939	<		hypothetical protein	Hypothetical, unclassified, unknown
as99	2121071	2121240	170	internal	PA1939	<		hypothetical protein	Hypothetical, unclassified, unknown
as100	2128729	2128826	98	internal	PA1944	>		hypothetical protein	Hypothetical, unclassified, unknown
as101	2314771	2314932	162	internal	PA2103	>		probable molybdopterin biosynthesis protein MoeB	Biosynthesis of cofactors, prosthetic groups and carriers
as102	2342397	2342560	164	start-span	PA2128	>	*cupA1*	fimbrial subunit CupA1	Motility & Attachment
as103	2400697	2400826	130	internal	PA2179	<		hypothetical protein	Hypothetical, unclassified, unknown
as104	2445134	2445251	118	internal	PA2222	<		hypothetical protein	Hypothetical, unclassified, unknown
as105	2550592	2550780	189	start-span	PA2313	>		hypothetical protein	Membrane proteins
as106	2616944	2617020	77	end-span	PA2366	>		uricase PuuD	Hypothetical, unclassified, unknown
as106	2616944	2617020	77	start-span	PA2367	>		hypothetical protein	Hypothetical, unclassified, unknown
as107	2623295	2623503	209	internal	PA2372	>		hypothetical protein	Hypothetical, unclassified, unknown
as108	2701170	2701296	127	start-span	PA2418	>		hypothetical protein	Hypothetical, unclassified, unknown
as109	2705716	2705781	66	start-span	PA2422	>		hypothetical protein	Hypothetical, unclassified, unknown
as110	2756154	2756267	114	end-span	PA2455	>		hypothetical protein	Hypothetical, unclassified, unknown
as111	2756364	2756944	581	end-span	PA2456	>		hypothetical protein	Hypothetical, unclassified, unknown
as111	2756364	2756944	581	start-span	PA2457	>		hypothetical protein	Hypothetical, unclassified, unknown
as112	2759171	2759360	190	end-span	PA2458	>		hypothetical protein	Hypothetical, unclassified, unknown
as113	2802778	2802933	156	internal	PA2484	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as114	2818927	2819159	233	end-span	PA2502	<		hypothetical protein	Hypothetical, unclassified, unknown
as115	2902292	2902391	100	internal	PA2567	>		hypothetical protein	Hypothetical, unclassified, unknown
as116	2905849	2905976	128	end-span	PA2570.1	<		tRNA-Leu	Non-coding RNA gene
as117	2946138	2946205	68	internal	PA2603	>		tRNA-Ser	Putative enzymes
as118	2959226	2959347	122	start-span	PA2615	<	*ftsK*	cell division protein FtsK	Cell division
as119	2959629	2959767	139	internal	PA2616	>	*trxB1*	thioredoxin reductase 1	Nucleotide biosynthesis and metabolism
as120	2982707	2982848	142	start-span	PA2637	>	*nuoA*	NADH dehydrogenase I chain A	Energy metabolism
as121	3005381	3005497	117	internal	PA2654	>		probable chemotaxis transducer	Adaptation, Protection; Chemotaxis
as122	3013744	3013890	147	start-span	PA2664	<	*fhp*	flavohemoprotein	Energy metabolism
as123	3015850	3015969	120	end-span	PA2666	>		probable 6-pyruvoyl tetrahydrobiopterin synthase	Biosynthesis of cofactors, prosthetic groups and carriers
as124	3016813	3016931	119	start-span	PA2668	<		hypothetical protein	Hypothetical, unclassified, unknown
as125	3060103	3060257	155	start-span	PA2705	<		hypothetical protein	Hypothetical, unclassified, unknown
as126	3110332	3110396	65	internal	PA2748	<		probable methionine aminopeptidase	Translation, post-translational modification, degradation
as127	3110413	3110473	61	internal	PA2748	<		probable methionine aminopeptidase	Translation, post-translational modification, degradation
as128	3119347	3119453	107	end-span	PA2759	<		hypothetical protein	Hypothetical, unclassified, unknown
as129	3120806	3120915	110	internal	PA2760	>	*OprQ*	oprQ	
as130	3167934	3168030	97	start-span	PA2814	<		hypothetical protein	Hypothetical, unclassified, unknown
as131	3173549	3173835	287	end-span	PA2819	>		tRNA-Glu	Hypothetical, unclassified, unknown
as132	3312901	3313118	218	internal	PA2953	>		electron transfer flavoprotein-ubiquinone oxidoreductase	Energy metabolism
as133	3357890	3357996	107	internal	PA3000	<	*aroP1*	aromatic amino acid transport protein AroP1	Transport of small molecules
as134	3395484	3395684	201	internal	PA3032	<	*snr1*	cytochrome c Snr1	Energy metabolism
as135	3414435	3414491	57	internal	PA3049	>	*rmf*	ribosome modulation factor	Translation, post-translational modification, degradation
as136	3508613	3508673	61	internal	PA3125	>		hypothetical protein	Hypothetical, unclassified, unknown
as137	3517655	3517722	68	internal	PA3135	<		probable transcriptional regulator	Transcriptional regulators
as138	3530607	3530794	188	internal	PA3147	<	*wbpJ*	probable glycosyl transferase WbpJ	Putative enzymes; Cell wall / LPS / capsule
as139	3534247	3534324	78	internal	PA3150	<	*wbpG*	LPS biosynthesis protein WbpG	Cell wall / LPS / capsule
as140	3539090	3539161	72	end-span	PA3155	<	*wbpE*	UDP-2-acetamido-2-dideoxy-d-ribo-hex-3-uluronic acid transaminase, wbpE	Putative enzymes; Cell wall / LPS / capsule
as140	3539090	3539161	72	start-span	PA3154	<	*wzy*	B-band O-antigen polymerase	Membrane proteins; Cell wall / LPS / capsule
as141	3540497	3540593	97	internal	PA3156	<	*wbpD*	UDP-2-acetamido-3-amino-2,3-dideoxy-d-glucuronic acid N-acetyltransferase, WbpD	Putative enzymes; Cell wall / LPS / capsule
as142	3540908	3541027	120	internal	PA3157	<		probable acetyltransferase	Putative enzymes; Cell wall / LPS / capsule
as143	3542803	3542997	195	internal	PA3158	<	*wbpB*	UDP-2-acetamido-2-deoxy-d-glucuronic acid 3-dehydrogenase, WbpB	Putative enzymes; Cell wall / LPS / capsule
as144	3592018	3592128	111	end-span	PA3200	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as144	3592018	3592128	111	start-span	PA3199	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as145	3593054	3593188	135	start-span	PA3201	>		conserved hypothetical protein	Membrane proteins
as146	3612043	3612282	240	internal	PA3224	<		hypothetical protein	Hypothetical, unclassified, unknown
as147	3625886	3625995	110	internal	PA3238	>		hypothetical protein	Hypothetical, unclassified, unknown
as148	3626660	3626776	117	internal	PA3239	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as149	3631510	3631629	120	end-span	PA3243	>	*minC*	cell division inhibitor MinC	Cell division
as149	3631510	3631629	120	start-span	PA3244	>	*minD*	cell division inhibitor MinD	Cell division
as150	3697506	3697616	111	internal	PA3300	<	*fadD2*	long-chain-fatty-acid--CoA ligase	Fatty acid and phospholipid metabolism
as151	3752069	3752155	87	internal	PA3341	<		probable transcriptional regulator	Transcriptional regulators
as152	3866017	3866206	190	internal	PA3459	>		probable glutamine amidotransferase	Amino acid biosynthesis and metabolism
as153	3889807	3889926	120	end-span	PA3477	<	*rhlR*	transcriptional regulator RhlR	Adaptation, Protection; Transcriptional regulators
as154	3944227	3944342	116	internal	PA3524	>	*gloA1*	lactoylglutathione lyase	Central intermediary metabolism
as155	3999082	3999164	83	end-span	PA3567	>		probable oxidoreductase	Putative enzymes
as156	4014211	4014326	116	internal	PA3582	>	*glpK*	glycerol kinase	Central intermediary metabolism
as157	4048252	4048377	126	start-span	PA3613	<		hypothetical protein	Hypothetical, unclassified, unknown
as158	4070600	4070735	136	internal	PA3636	<	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	Carbon compound catabolism; Cell wall / LPS / capsule
as159	4087208	4087269	62	internal	PA3648	<	*opr86*	outer membrane protein Opr86	Membrane proteins; Transport of small molecules
as160	4092478	4092590	113	internal	PA3654	<	*pyrH*	uridylate kinase	Nucleotide biosynthesis and metabolism
as161	4110202	4110321	120	internal	PA3670	<		hypothetical protein	Hypothetical, unclassified, unknown
as162	4156819	4156946	128	internal	PA3711	>		probable transcriptional regulator	Transcriptional regulators
as163	4166790	4167046	257	internal	PA3721	>	*nalC*	NalC	Transcriptional regulators; Antibiotic resistance and susceptibility
as164	4185534	4185601	68	internal	PA3734	>		hypothetical protein	Hypothetical, unclassified, unknown
as165	4242856	4242966	111	internal	PA3783	>		hypothetical protein	Hypothetical, unclassified, unknown
as166	4311907	4312074	168	end-span	PA3851	<		hypothetical protein	Hypothetical, unclassified, unknown
as167	4327946	4328181	236	internal	PA3866	<		pyocin protein	Adaptation, Protection; Secreted Factors (toxins, enzymes, alginate)
as168	4328357	4328518	162	internal	PA3866	<		pyocin protein	Adaptation, Protection; Secreted Factors (toxins, enzymes, alginate)
as169	4329401	4329645	245	internal	PA3866	<		pyocin protein	Adaptation, Protection; Secreted Factors (toxins, enzymes, alginate)
as170	4330879	4330989	111	end-span	PA3867	>		probable DNA invertase	DNA replication, recombination, modification and repair
as171	4471922	4472053	132	internal	PA3991	<		hypothetical protein	Hypothetical, unclassified, unknown
as172	4616918	4616968	51	internal	PA4128	>		conserved hypothetical protein	Putative enzymes
as173	4619821	4619903	83	internal	PA4131	>		probable iron-sulfur protein	Putative enzymes
as174	4666087	4666202	116	internal	PA4168	>	*fpvB*	second ferric pyoverdine receptor FpvB	Transport of small molecules
as175	4711769	4711861	93	internal	PA4208	>	*opmD*	probable outer membrane protein precursor	Membrane proteins; Transport of small molecules
as176	4759525	4759686	162	end-span	PA4247	<	*rplR*	50S ribosomal protein L18	Translation, post-translational modification, degradation
as176	4759525	4759686	162	start-span	PA4246	<	*rpsE*	30S ribosomal protein S5	Translation, post-translational modification, degradation
as177	4760853	4760902	50	start-span	PA4249	<	*rpsH*	30S ribosomal protein S8	Translation, post-translational modification, degradation
as178	4761521	4761592	72	internal	PA4251	<	*rplE*	50S ribosomal protein L5	Translation, post-translational modification, degradation
as179	4763739	4763790	52	internal	PA4257	<	*rpsC*	30S ribosomal protein S3	Translation, post-translational modification, degradation
as180	4764919	4765009	91	internal	PA4260	<	*rplB*	50S ribosomal protein L2	Translation, post-translational modification, degradation
as181	4773392	4773503	112	internal	PA4269	<	*rpoC*	DNA-directed RNA polymerase beta* chain	Transcription, RNA processing and degradation

as182	4780416	4780494	79	internal	PA4270	<	*rpoB*	P26	Transcription, RNA processing and degradation
as183	4781848	4781947	100	internal	PA4272.1	>	*P27*	P27	Non-coding RNA gene
as184	4842427	4842583	157	end-span	PA4313	>		hypothetical protein	Hypothetical, unclassified, unknown
as185	4919101	4919236	136	internal	PA4389	>		probable short-chain dehydrogenase	Putative enzymes; Amino acid biosynthesis and metabolism
as186	4925623	4925733	111	internal	PA4395	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as187	4982939	4983086	148	internal	PA4448	<	*hisD*	histidinol dehydrogenase	Amino acid biosynthesis and metabolism
as188	4991277	4991330	54	internal	PA4459	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as189	5002347	5002404	58	internal	PA4473	>		hypothetical protein	Hypothetical, unclassified, unknown
as190	5018832	5019000	169	internal	PA4488	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as191	5037299	5037430	132	internal	PA4500	>		probable binding protein component of ABC transporter	Transport of small molecules
as192	5049739	5049879	141	start-span	PA4512	>	*lpxO1*	lipopolysaccharide biosynthetic protein LpxO1	Putative enzymes; Cell wall / LPS / capsule
as193	5055080	5055129	50	internal	PA4514	<		probable outer membrane receptor for iron transport	Transport of small molecules
as194	5055132	5055185	54	internal	PA4514	<		probable outer membrane receptor for iron transport	Transport of small molecules
as195	5071575	5071963	389	internal	PA4527	>	*pilC*	still frameshift type 4 fimbrial biogenesis protein PilC	Motility & Attachment
as196	5071970	5072154	185	internal	PA4527	>	*pilC*	still frameshift type 4 fimbrial biogenesis protein PilC	Motility & Attachment
as197	5101107	5101200	94	internal	PA4554	>	*pilY1*	type 4 fimbrial biogenesis protein PilY1	Motility & Attachment
as198	5101511	5101611	101	internal	PA4554	>	*pilY1*	type 4 fimbrial biogenesis protein PilY1	Motility & Attachment
as199	5141781	5141969	189	end-span	PA4591	<		hypothetical protein	Hypothetical, unclassified, unknown
as200	5148538	5148668	131	start-span	PA4595	<		probable ATP-binding component of ABC transporter	Transport of small molecules
as201	5274032	5274193	162	internal	PA4696	<	*ilvI*	acetolactate synthase large subunit	Biosynthesis of cofactors, prosthetic groups and carriers; Amino acid biosynthesis and metabolism
as202	5296232	5296293	62	internal	PA4717	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as203	5296311	5296538	228	internal	PA4717	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as204	5301771	5301838	68	internal	PA4722	<		probable aminotransferase	Putative enzymes
as205	5356355	5356495	141	internal	PA4770	>	*lldP*	L-lactate permease	Transport of small molecules
as206	5375454	5375529	76	end-span	PA4786	>		probable short-chain dehydrogenase	Putative enzymes
as207	5380813	5380919	107	internal	PA4793	>		hypothetical protein	Hypothetical, unclassified, unknown
as208	5382370	5382497	128	end-span	PA4796	>		hypothetical protein	Hypothetical, unclassified, unknown
as209	5408036	5408137	102	internal	PA4818	<		conserved hypothetical protein	Membrane proteins
as210	5473394	5473504	111	internal	PA4877	<		hypothetical protein	Hypothetical, unclassified, unknown
as211	5522339	5522410	72	start-span	PA4923	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as212	5532791	5532916	126	internal	PA4929	>		hypothetical protein	Membrane proteins
as213	5541691	5541800	110	end-span	PA4937.1	>		tRNA-Leu	Non-coding RNA gene
as214	5545002	5545131	130	internal	PA4941	<	*hflC*	protease subunit HflC	Translation, post-translational modification, degradation; Cell division
as215	5559871	5559952	82	start-span	PA4954	<	*motA*	chemotaxis protein MotA	Adaptation, Protection; Chemotaxis
as216	5580862	5581109	248	start-span	PA4972	>		hypothetical protein	Hypothetical, unclassified, unknown
as217	5584935	5585163	229	internal	PA4974	>		probable outer membrane protein precursor	Protein secretion/export apparatus
as218	5585580	5585652	73	end-span	PA4975	<		NAD(P)H quinone oxidoreductase	Energy metabolism
as219	5602759	5602870	112	internal	PA4986	>		probable oxidoreductase	Putative enzymes
as220	5617657	5617793	137	internal	PA5000	<	*wapR*	alpha-1,3-rhamnosyltransferase WapR	Putative enzymes; Cell wall / LPS / capsule
as221	5618584	5618634	51	internal	PA5001	<		hypothetical protein	Hypothetical, unclassified, unknown
as222	5655041	5655112	72	internal	PA5024	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as223	5678557	5678652	96	internal	PA5042	<	*pilO*	type 4 fimbrial biogenesis protein PilO	Motility & Attachment
asponA****	5680819	5681167	349	start-span	PA5045	>	*ponA*	penicillin-binding protein 1A	Cell wall / LPS / capsule
as225	5684803	5684910	108	end-span	PA5047	<		hypothetical protein	Hypothetical, unclassified, unknown
as226	5699012	5699062	51	internal	PA5058	>	*phaC2*	poly(3-hydroxyalkanoic acid) synthase 2	Central intermediary metabolism
as227	5720268	5720351	84	start-span	PA5080	<		prolyl aminopeptidase	Translation, post-translational modification, degradation
as228	5725557	5725795	239	internal	PA5087	<		hypothetical protein	Hypothetical, unclassified, unknown
as229	5739066	5739203	138	internal	PA5097	<		probable amino acid permease	Membrane proteins; Transport of small molecules
as230	5785777	5785905	129	internal	PA5136	>		hypothetical protein	Hypothetical, unclassified, unknown
as231	5799079	5799208	130	internal	PA5150	<		probable short-chain dehydrogenase	Putative enzymes
as232	5817430	5817493	64	internal	PA5167	>		probable c4-dicarboxylate-binding protein	Membrane proteins; Transport of small molecules
as233	5824524	5824599	76	internal	PA5172	>	*arcB*	ornithine carbamoyltransferase, catabolic	Amino acid biosynthesis and metabolism
as234	5844451	5844527	77	end-span	PA5192	<	*pckA*	phosphoenolpyruvate carboxykinase	Carbon compound catabolism; Energy metabolism
as235	5863452	5863719	268	start-span	PA5208	<		conserved hypothetical protein	Hypothetical, unclassified, unknown
as236	5868209	5868408	200	internal	PA5213	<	*gcvP1*	glycine cleavage system protein P1	Central intermediary metabolism; Amino acid biosynthesis and metabolism
as237	5896925	5897003	79	internal	PA5238	<		probable O-antigen acetylase	Putative enzymes; Membrane proteins; Cell wall / LPS / capsule
as238	5897008	5897057	50	internal	PA5238	<		probable O-antigen acetylase	Putative enzymes; Membrane proteins; Cell wall / LPS / capsule
as239	5926819	5926963	145	internal	PA5264	<		hypothetical protein	Hypothetical, unclassified, unknown; Membrane proteins
as240	6059529	6059600	72	internal	PA5380	>	*gbdR*	GbdR	Transcriptional regulators
as241	6090655	6090800	146	start-span	PA5412	<		hypothetical protein	Hypothetical, unclassified, unknown
as242	6135711	6135789	79	internal	PA5446	>		hypothetical protein	Hypothetical, unclassified, unknown
as243	6141133	6141239	107	internal	PA5451	<	*wzm*	membrane subunit of A-band LPS efflux transporter	Cell wall / LPS / capsule; Membrane proteins; Transport of small molecules
as244	6157855	6157964	110	internal	PA5468	>		probable citrate transporter	Transport of small molecules
as245	6158032	6158203	172	end-span	PA5468	>		probable citrate transporter	Transport of small molecules
as245	6158032	6158203	172	start-span	PA5469	>		conserved hypothetical protein	Membrane proteins
as246	6171684	6171915	232	start-span	PA5480	>		hypothetical protein	Hypothetical, unclassified, unknown
as247	6183216	6183389	174	internal	PA5492	>		conserved hypothetical protein	Hypothetical, unclassified, unknown
as248	6218749	6219078	330	end-span	PA5525	>		probable transcriptional regulator	Transcriptional regulators
as249	6245830	6245973	144	internal	PA5551	<		hypothetical protein	Hypothetical, unclassified, unknown
as250	6247796	6247909	114	end-span	PA5553	<	*atpC*	ATP synthase epsilon chain	Energy metabolism

* Coordinates and length of the sRNAs predicted by RNA-Seq data.

**The antisense sRNAs were categorize depending on their position with respect to the opposite gene as “start-span” (the asRNA spans over the start of the antisense gene), “internal” (the asRNA start and end within the antisense gene), or “end-span” (the asRNA spans over the end of the gene).

***Some asRNAs overlap with 2 genes.

****Coordinates validated by 5´and 3´-RACE experiments.

Supplementary Table 3. Antisense sRNAs differentially expressed.

Name	Start*	End*	Length*	Direction asRNA	Fold change**
asRNAs differentially expressed during AZI (10xMIC) treatment					
as165	4242856	4242966	111	<	0.12
as210	5473394	5473504	111	>	0.21
as94	2027457	2027571	115	<	0.26
as230	5785777	5785905	129	<	0.32
as77	1640773	1640948	176	>	0.38
as100	2128729	2128826	98	<	0.39
as42	977508	977658	151	>	0.42
as183	4781848	4781947	100	<	2.55
as130	3167934	3168030	97	>	2.66
as99	2121071	2121240	170	>	2.69
as123	3015850	3015969	120	<	2.90
as231	5799079	5799208	130	>	2.94
as115	2902292	2902391	100	<	3.30
as111	2756364	2756944	581	<	3.34
as112	2759171	2759360	190	<	3.88
as117	2946138	2946205	68	<	4.87
as244	6157855	6157964	110	<	11.33
as184	4842427	4842583	157	<	27.49
asRNAs differentially expressed during AZT (3xMIC) treatment					
as6	182516	182617	102	<	0.09
as48	1060351	1060676	326	>	0.26
as104	2445134	2445251	118	>	0.32
as19	448929	449284	356	>	0.33
as119	2959629	2959767	139	<	0.36
as78	1698424	1698534	111	>	0.41
as120	2982707	2982848	142	<	2.06
as29	701233	701489	257	<	2.07
as163	4166790	4167046	257	<	2.20
as10	288493	288639	147	>	2.30
as200	5148538	5148668	131	>	2.33
as228	5725557	5725795	239	>	2.34
as195	5071575	5071963	389	<	2.51
asdnaA****	1320	1784	465	<	2.79
as146	3612043	3612282	240	>	2.82
as51	1087094	1087222	129	>	2.85
asponA****	5680819	5681167	349	<	2.94
as205	5356355	5356495	141	<	3.13
as162	4156819	4156946	128	<	3.24
as45	1027896	1028083	188	<	3.56
as135	3414435	3414491	57	<	4.21
as181	4773392	4773503	112	>	5.50
as221	5618584	5618634	51	>	6.08
as244	6157855	6157964	110	<	6.37
as32	749460	749562	103	<	6.84
as39	921673	921840	168	>	20.82
as152	3866017	3866206	190	<	21.77
as128	3119347	3119453	107	>	22.06
asRNAs differentially expressed during CEF (3xMIC) treatment					
as6	182516	182617	102	<	0.08
as75	1588434	1588549	116	<	0.09

as210	5473394	5473504	111	>	0.12
as48	1060351	1060676	326	>	0.14
as83	1838134	1838246	113	<	0.16
as119	2959629	2959767	139	<	0.17
as19	448929	449284	356	>	0.20
as43	990194	990309	116	<	0.35
as138	3530607	3530794	188	>	0.37
as94	2027457	2027571	115	<	0.39
as235	5863452	5863719	268	>	0.43
as146	3612043	3612282	240	>	2.10
as29	701233	701489	257	<	2.17
as24	646779	646892	114	>	2.21
as200	5148538	5148668	131	>	2.22
as45	1027896	1028083	188	<	2.32
as192	5049739	5049879	141	<	2.45
as51	1087094	1087222	129	>	3.15
as4	35866	35974	109	>	4.02
as16	353325	353458	134	>	4.93
as213	5541691	5541800	110	<	7.18
as145	3593054	3593188	135	<	7.72
as129	3120806	3120915	110	<	12.58
as39	921673	921840	168	>	24.03
as56	1141250	1141364	115	<	32.30
asRNAs differentially expressed during CIP (3xMIC) treatment					
as6	182516	182617	102	<	0.02
as92	2012282	2012454	173	>	0.05
as64	1484863	1485037	175	<	0.07
as169	4329401	4329645	245	>	0.07
as104	2445134	2445251	118	>	0.10
as157	4048252	4048377	126	>	0.11
as183	4781848	4781947	100	<	0.11
as131	3173549	3173835	287	<	0.13
as110	2756154	2756267	114	<	0.15
as23	557312	557569	258	>	0.17
as111	2756364	2756944	581	<	0.19
as231	5799079	5799208	130	>	0.20
as29	701233	701489	257	<	0.20
as21	544658	544778	121	<	0.21
as77	1640773	1640948	176	>	0.23
as19	448929	449284	356	>	0.23
as225	5684803	5684910	108	>	0.24
as165	4242856	4242966	111	<	0.24
as31	748895	749047	153	<	0.27
as12	291428	291543	116	>	0.31
as101	2314771	2314932	162	<	0.34
as48	1060351	1060676	326	>	0.34
as55	1135776	1135875	100	<	0.36
as38	898637	898918	282	>	0.37
as43	990194	990309	116	<	0.38
as94	2027457	2027571	115	<	0.40
as98	2120607	2120921	315	>	0.43
as138	3530607	3530794	188	>	0.47
as235	5863452	5863719	268	>	0.49
as9	263722	263909	188	<	0.49
as100	2128729	2128826	98	<	2.05
asponA****	5680819	5681167	349	<	2.06

as179	4763739	4763790	52	>	2.22
as208	5382370	5382497	128	<	2.48
as52	1098349	1098557	209	>	2.80
as26	659114	659270	157	>	2.84
as160	4092478	4092590	113	>	2.97
as163	4166790	4167046	257	<	3.05
as7	184720	184861	142	<	3.26
as24	646779	646892	114	>	3.35
as45	1027896	1028083	188	<	3.54
as181	4773392	4773503	112	>	3.69
as32	749460	749562	103	<	3.77
as115	2902292	2902391	100	<	4.45
asdnaA****	1320	1784	465	<	4.67
as136	3508613	3508673	61	<	6.06
as4	35866	35974	109	>	6.41
as152	3866017	3866206	190	<	6.96
as170	4330879	4330989	111	<	10.94
as16	353325	353458	134	>	12.17
as128	3119347	3119453	107	>	16.61
as122	3013744	3013890	147	>	22.25
as129	3120806	3120915	110	<	24.88
as213	5541691	5541800	110	<	30.64
asRNAs differentially expressed during CO (3xMIC) treatment					
as6	182516	182617	102	<	0.02
as37	817964	818095	132	<	0.10
as19	448929	449284	356	>	0.11
as230	5785777	5785905	129	<	0.12
as201	5274032	5274193	162	>	0.14
as131	3173549	3173835	287	<	0.14
as64	1484863	1485037	175	<	0.18
as104	2445134	2445251	118	>	0.18
as235	5863452	5863719	268	>	0.19
as48	1060351	1060676	326	>	0.20
as83	1838134	1838246	113	<	0.20
as92	2012282	2012454	173	>	0.20
as12	291428	291543	116	>	0.21
as206	5375454	5375529	76	<	0.30
as9	263722	263909	188	<	0.30
as203	5296311	5296538	228	<	0.32
as42	977508	977658	151	>	0.32
asponA****	5680819	5681167	349	<	0.32
as225	5684803	5684910	108	>	0.33
as112	2759171	2759360	190	<	0.33
as11	290529	290665	137	>	0.34
as55	1135776	1135875	100	<	0.34
as44	1026460	1026563	104	>	0.36
as61	1400645	1400720	76	>	0.36
as169	4329401	4329645	245	>	0.39
as82	1790530	1790650	121	>	0.42
as63	1474290	1474393	104	<	0.43
as81	1785747	1785881	135	>	0.45
as146	3612043	3612282	240	>	2.25
as119	2959629	2959767	139	<	2.30
as192	5049739	5049879	141	<	2.37
as90	1954335	1954415	81	<	2.72
as120	2982707	2982848	142	<	2.95

as78	1698424	1698534	111	>	3.04
as16	353325	353458	134	>	3.41
as221	5618584	5618634	51	>	3.87
as247	6183216	6183389	174	<	4.48
as205	5356355	5356495	141	<	4.96
as162	4156819	4156946	128	<	5.35
as57	1207482	1207531	50	<	5.44
as121	3005381	3005497	117	<	6.22
as25	654775	654912	138	>	6.25
as36	798624	798854	231	>	6.86
as20	521726	521838	113	>	36.01
asRNAs differentially expressed during H2O2 (1 mM) treatment					
as6	182516	182617	102	<	0.06
as48	1060351	1060676	326	>	0.35
as104	2445134	2445251	118	>	0.45
as19	448929	449284	356	>	0.45
as84	1911985	1912118	134	>	2.05
asdnaA****	1320	1784	465	<	2.11
as29	701233	701489	257	<	2.28
as146	3612043	3612282	240	>	2.96
as174	4666087	4666202	116	<	3.31
as205	5356355	5356495	141	<	3.89
as143	3542803	3542997	195	>	4.70
as244	6157855	6157964	110	<	5.40
as120	2982707	2982848	142	<	5.79
as221	5618584	5618634	51	>	6.24
as145	3593054	3593188	135	<	6.51
as20	521726	521838	113	>	7.47
as39	921673	921840	168	>	9.95
asRNAs differentially expressed during MP (3xMIC) treatment					
as6	182516	182617	102	<	0.05
as19	448929	449284	356	>	0.09
as201	5274032	5274193	162	>	0.09
as77	1640773	1640948	176	>	0.09
as104	2445134	2445251	118	>	0.10
as235	5863452	5863719	268	>	0.14
as12	291428	291543	116	>	0.15
as131	3173549	3173835	287	<	0.16
as48	1060351	1060676	326	>	0.16
as55	1135776	1135875	100	<	0.19
as230	5785777	5785905	129	<	0.22
as138	3530607	3530794	188	>	0.26
as63	1474290	1474393	104	<	0.27
as169	4329401	4329645	245	>	0.27
as225	5684803	5684910	108	>	0.31
as220	5617657	5617793	137	>	0.33
as203	5296311	5296538	228	<	0.41
as31	748895	749047	153	<	0.44
as94	2027457	2027571	115	<	0.45
as23	557312	557569	258	>	0.45
asponA****	5680819	5681167	349	<	0.48
as81	1785747	1785881	135	>	2.20
as146	3612043	3612282	240	>	2.62
as200	5148538	5148668	131	>	2.82
as16	353325	353458	134	>	3.57
as163	4166790	4167046	257	<	3.92

as248	6218749	6219078	330	<	4.08
as29	701233	701489	257	<	4.51
as162	4156819	4156946	128	<	5.46
as130	3167934	3168030	97	>	14.81
as62	1455941	1456046	106	>	15.67
as213	5541691	5541800	110	<	15.76
as236	5868209	5868408	200	>	21.07
asRNAs differentially expressed during NaCl (0.5 M) treatment					
as6	182516	182617	102	<	0.12
as12	291428	291543	116	>	0.16
as19	448929	449284	356	>	0.20
as48	1060351	1060676	326	>	0.27
as55	1135776	1135875	100	<	0.33
as235	5863452	5863719	268	>	0.36
as74	1558500	1558752	253	<	0.41
as84	1911985	1912118	134	>	2.55
as214	5545002	5545131	130	>	2.71
as116	2905849	2905976	128	>	2.78
as36	798624	798854	231	>	3.58
as236	5868209	5868408	200	>	4.13
as204	5301771	5301838	68	>	6.66
as130	3167934	3168030	97	>	11.83
as62	1455941	1456046	106	>	12.37
asRNAs differentially expressed during PP (3xMIC) treatment					
as6	182516	182617	102	<	0.07
as104	2445134	2445251	118	>	0.09
as48	1060351	1060676	326	>	0.25
as94	2027457	2027571	115	<	0.29
as19	448929	449284	356	>	0.34
as138	3530607	3530794	188	>	0.39
as43	990194	990309	116	<	0.40
as92	2012282	2012454	173	>	0.42
as63	1474290	1474393	104	<	0.48
as55	1135776	1135875	100	<	0.50
as212	5532791	5532916	126	<	2.11
as181	4773392	4773503	112	>	2.31
as44	1026460	1026563	104	>	2.36
as162	4156819	4156946	128	<	2.47
as96	2113618	2113734	117	<	2.71
as143	3542803	3542997	195	>	2.79
asponA****	5680819	5681167	349	<	2.83
as146	3612043	3612282	240	>	2.95
as91	1954420	1954469	50	<	3.14
as170	4330879	4330989	111	<	4.03
as221	5618584	5618634	51	>	4.17
as205	5356355	5356495	141	<	4.22
as16	353325	353458	134	>	4.25
as244	6157855	6157964	110	<	6.49
as129	3120806	3120915	110	<	21.17
asRNAs differentially expressed during TET (3xMIC) treatment					
as92	2012282	2012454	173	>	0.02
as78	1698424	1698534	111	>	0.04
as19	448929	449284	356	>	0.05
as165	4242856	4242966	111	<	0.07
as21	544658	544778	121	<	0.08
as12	291428	291543	116	>	0.10

as6	182516	182617	102	<	0.13
as26	659114	659270	157	>	0.14
as61	1400645	1400720	76	>	0.15
as217	5584935	5585163	229	<	0.15
as94	2027457	2027571	115	<	0.16
asponA****	5680819	5681167	349	<	0.18
as31	748895	749047	153	<	0.20
as120	2982707	2982848	142	<	0.24
as42	977508	977658	151	>	0.25
as157	4048252	4048377	126	>	0.26
as53	1119615	1119886	272	>	0.26
as29	701233	701489	257	<	0.28
as9	263722	263909	188	<	0.31
as23	557312	557569	258	>	0.33
as192	5049739	5049879	141	<	0.36
as250	6247796	6247909	114	>	0.40
as144	3592018	3592128	111	>	0.43
as206	5375454	5375529	76	<	0.47
as67	1499690	1499742	53	<	0.49
as74	1558500	1558752	253	<	2.30
as60	1394817	1395002	186	>	2.40
as11	290529	290665	137	>	2.43
as154	3944227	3944342	116	<	2.98
as111	2756364	2756944	581	<	3.03
as162	4156819	4156946	128	<	3.51
as104	2445134	2445251	118	>	3.72
as163	4166790	4167046	257	<	4.18
as239	5926819	5926963	145	>	4.37
as142	3540908	3541027	120	>	4.83
as203	5296311	5296538	228	<	5.44
as54	1135102	1135286	185	<	5.69
as236	5868209	5868408	200	>	6.63
as241	6090655	6090800	146	>	6.76
as30	718785	718934	150	>	6.97
as136	3508613	3508673	61	<	7.51
as210	5473394	5473504	111	>	7.66
as99	2121071	2121240	170	>	7.79
as107	2623295	2623503	209	<	9.26
as25	654775	654912	138	>	10.98
as16	353325	353458	134	>	11.47
as64	1484863	1485037	175	<	11.74
as216	5580862	5581109	248	<	12.15
as244	6157855	6157964	110	<	14.57
as112	2759171	2759360	190	<	14.91
as167	4327946	4328181	236	>	16.07
as88	1949036	1949142	107	>	19.62
as117	2946138	2946205	68	<	20.08
as49	1066483	1066610	128	>	20.16
as20	521726	521838	113	>	21.57
as76	1639661	1639805	145	>	24.33
as50	1073865	1074007	143	<	25.41
as134	3395484	3395684	201	>	26.97
as47	1056103	1056248	146	<	27.11
as115	2902292	2902391	100	<	42.28
as247	6183216	6183389	174	<	43.19

asRNAs differentially expressed during TOB (3xMIC) treatment

as92	2012282	2012454	173	>	0.07
as165	4242856	4242966	111	<	0.08
as42	977508	977658	151	>	0.11
as120	2982707	2982848	142	<	0.13
as6	182516	182617	102	<	0.14
as78	1698424	1698534	111	>	0.16
as19	448929	449284	356	>	0.17
asponA****	5680819	5681167	349	<	0.18
as119	2959629	2959767	139	<	0.21
as74	1558500	1558752	253	<	0.25
as12	291428	291543	116	>	0.36
as63	1474290	1474393	104	<	0.37
as26	659114	659270	157	>	0.40
as206	5375454	5375529	76	<	0.40
as183	4781848	4781947	100	<	0.41
as21	544658	544778	121	<	0.45
as200	5148538	5148668	131	>	0.47
as31	748895	749047	153	<	0.48
as7	184720	184861	142	<	2.07
as214	5545002	5545131	130	>	2.31
as77	1640773	1640948	176	>	2.36
as195	5071575	5071963	389	<	2.40
as250	6247796	6247909	114	>	2.43
as39	921673	921840	168	>	2.46
as38	898637	898918	282	>	2.53
as146	3612043	3612282	240	>	2.55
as168	4328357	4328518	162	>	2.67
as203	5296311	5296538	228	<	2.79
as179	4763739	4763790	52	>	2.94
as176	4759525	4759686	162	>	3.38
as233	5824524	5824599	76	<	3.67
as107	2623295	2623503	209	<	3.75
as182	4780416	4780494	79	>	3.80
as247	6183216	6183389	174	<	3.82
as97	2113817	2113927	111	<	3.94
as115	2902292	2902391	100	<	4.15
as52	1098349	1098557	209	>	4.22
as45	1027896	1028083	188	<	4.34
as27	680736	680853	118	<	4.41
as50	1073865	1074007	143	<	4.57
as76	1639661	1639805	145	>	4.60
as95	2113335	2113442	108	<	5.62
as25	654775	654912	138	>	6.19
as138	3530607	3530794	188	>	6.37
as64	1484863	1485037	175	<	6.41
as99	2121071	2121240	170	>	7.55
as110	2756154	2756267	114	<	7.72
as62	1455941	1456046	106	>	7.76
as30	718785	718934	150	>	8.27
as236	5868209	5868408	200	>	9.24
as130	3167934	3168030	97	>	10.13
as96	2113618	2113734	117	<	11.00
as205	5356355	5356495	141	<	11.76
as93	2027111	2027193	83	<	14.45
as162	4156819	4156946	128	<	14.83
as16	353325	353458	134	>	18.51

as71	1504553	1504676	124	<	20.05
as245	6158032	6158203	172	<	20.79
as118	2959226	2959347	122	>	27.44
as70	1503843	1503937	95	<	30.29
as244	6157855	6157964	110	<	37.70
as136	3508613	3508673	61	<	41.81
as20	521726	521838	113	>	51.70
as36	798624	798854	231	>	104.80
asRNAs differentially expressed during early stationary phase					
as6	182516	182617	102	<	0.02
as231	5799079	5799208	130	>	0.11
as78	1698424	1698534	111	>	0.05
as201	5274032	5274193	162	>	0.05
as75	1588434	1588549	116	<	0.10
as161	4110202	4110321	120	>	0.11
as190	5018832	5019000	169	>	0.11
as44	1026460	1026563	104	>	0.13
as121	3005381	3005497	117	<	0.15
as230	5785777	5785905	129	<	0.16
as212	5532791	5532916	126	<	0.18
as203	5296311	5296538	228	<	0.19
as82	1790530	1790650	121	>	0.20
as48	1060351	1060676	326	>	0.22
as26	659114	659270	157	>	0.24
as19	448929	449284	356	>	0.26
as7	184720	184861	142	<	0.27
as104	2445134	2445251	118	>	0.30
as43	990194	990309	116	<	0.31
as165	4242856	4242966	111	<	0.32
as216	5580862	5581109	248	<	0.32
as210	5473394	5473504	111	>	0.32
as23	557312	557569	258	>	0.32
as74	1558500	1558752	253	<	0.35
as81	1785747	1785881	135	>	0.36
as21	544658	544778	121	<	0.36
as42	977508	977658	151	>	0.37
as138	3530607	3530794	188	>	0.44
as235	5863452	5863719	268	>	0.47
as144	3592018	3592128	111	>	0.48
as55	1135776	1135875	100	<	0.49
as94	2027457	2027571	115	<	0.49
as9	263722	263909	188	<	0.56
as53	1119615	1119886	272	>	0.57
as92	2012282	2012454	173	>	0.70
as31	748895	749047	153	<	0.90
as60	1394817	1395002	186	>	1.51
as29	701233	701489	257	<	1.70
as67	1499690	1499742	53	<	1.83
as214	5545002	5545131	130	>	1.89
as142	3540908	3541027	120	>	2.00
as143	3542803	3542997	195	>	2.09
as200	5148538	5148668	131	>	2.39
as111	2756364	2756944	581	<	2.77
as38	898637	898918	282	>	2.91
as84	1911985	1912118	134	>	3.18
as68	1502841	1502987	147	<	3.27

as204	5301771	5301838	68	>	3.64
as110	2756154	2756267	114	<	3.67
as228	5725557	5725795	239	>	3.91
as168	4328357	4328518	162	>	3.91
as157	4048252	4048377	126	>	4.10
as101	2314771	2314932	162	<	4.12
as194	5055132	5055185	54	>	4.31
as158	4070600	4070735	136	>	4.60
as169	4329401	4329645	245	>	5.21
as112	2759171	2759360	190	<	5.25
as197	5101107	5101200	94	<	5.35
as249	6245830	6245973	144	>	7.19
as99	2121071	2121240	170	>	7.48
as135	3414435	3414491	57	<	7.76
as173	4619821	4619903	83	<	7.97
as177	4760853	4760902	50	>	8.22
asdnaA****	1656	1813	158	<	8.70
as62	1455941	1456046	106	>	9.45
as238	5897008	5897057	50	>	11.15
as239	5926819	5926963	145	>	11.43
as195	5071575	5071963	389	<	12.18
as218	5585580	5585652	73	>	12.18
as148	3626660	3626776	117	<	12.90
as147	3625886	3625995	110	<	14.45
as221	5618584	5618634	51	>	14.92
as11	290529	290665	137	>	15.87
as182	4780416	4780494	79	>	16.64
as176	4759525	4759686	162	>	17.15
as34	784213	784303	91	<	18.57
as8	226915	226976	62	<	18.59
as151	3752069	3752155	87	>	20.20
as105	2550592	2550780	189	<	20.43
as196	5071970	5072154	185	<	21.03
as141	3540497	3540593	97	>	21.54
as232	5817430	5817493	64	<	22.70
as133	3357890	3357996	107	>	23.80
as186	4925623	4925733	111	>	23.83
as114	2818927	2819159	233	>	24.56
as149	3631510	3631629	120	<	25.52
as89	1953270	1953371	102	<	28.32
as87	1916046	1916172	127	<	28.79
as178	4761521	4761592	72	>	29.29
as223	5678557	5678652	96	>	30.13
as227	5720268	5720351	84	>	33.56
as199	5141781	5141969	189	>	34.59
as237	5896925	5897003	79	>	39.97
as145	3593054	3593188	135	<	42.93
as242	6135711	6135789	79	<	45.32
as109	2705716	2705781	66	<	50.81
as13	299094	299181	88	>	57.43
as56	1141250	1141364	115	<	66.44
as137	3517655	3517722	68	>	67.12
as229	5739066	5739203	138	>	72.49
as108	2701170	2701296	127	<	61.53
as122	3013744	3013890	147	>	55.57
as28	697609	697703	95	<	77.29

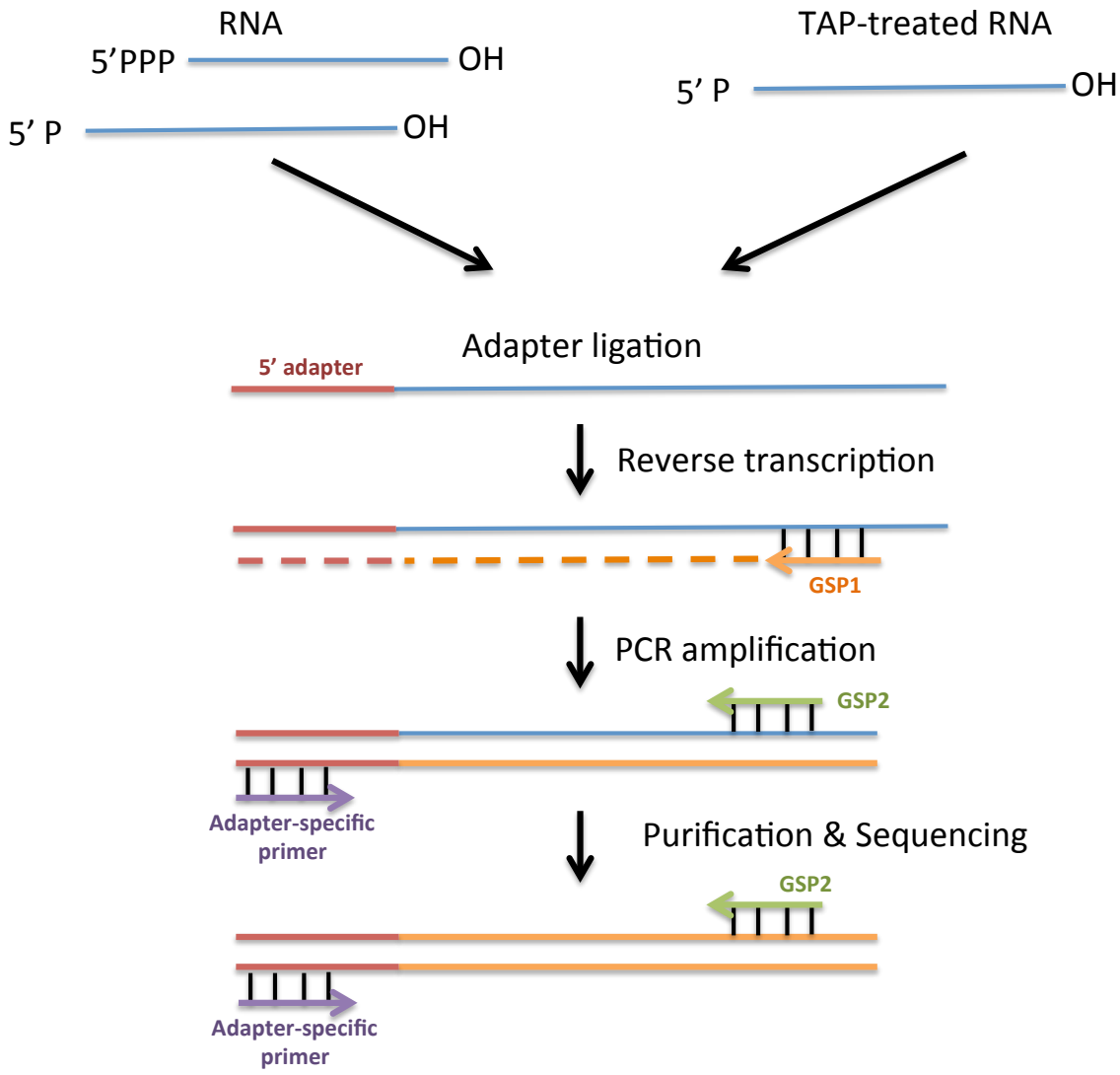
as164	4185534	4185601	68	<	70.65
as3	15121	15215	95	<	128.92
as189	5002347	5002404	58	<	71.99
as226	5699012	5699062	51	<	77.00
as59	1323807	1323918	112	>	70.06
as159	4087208	4087269	62	>	73.52
as222	5655041	5655112	72	<	79.82
as172	4616918	4616968	51	<	92.70

* Coordinates and length of the sRNAs predicted by RNA-Seq data.

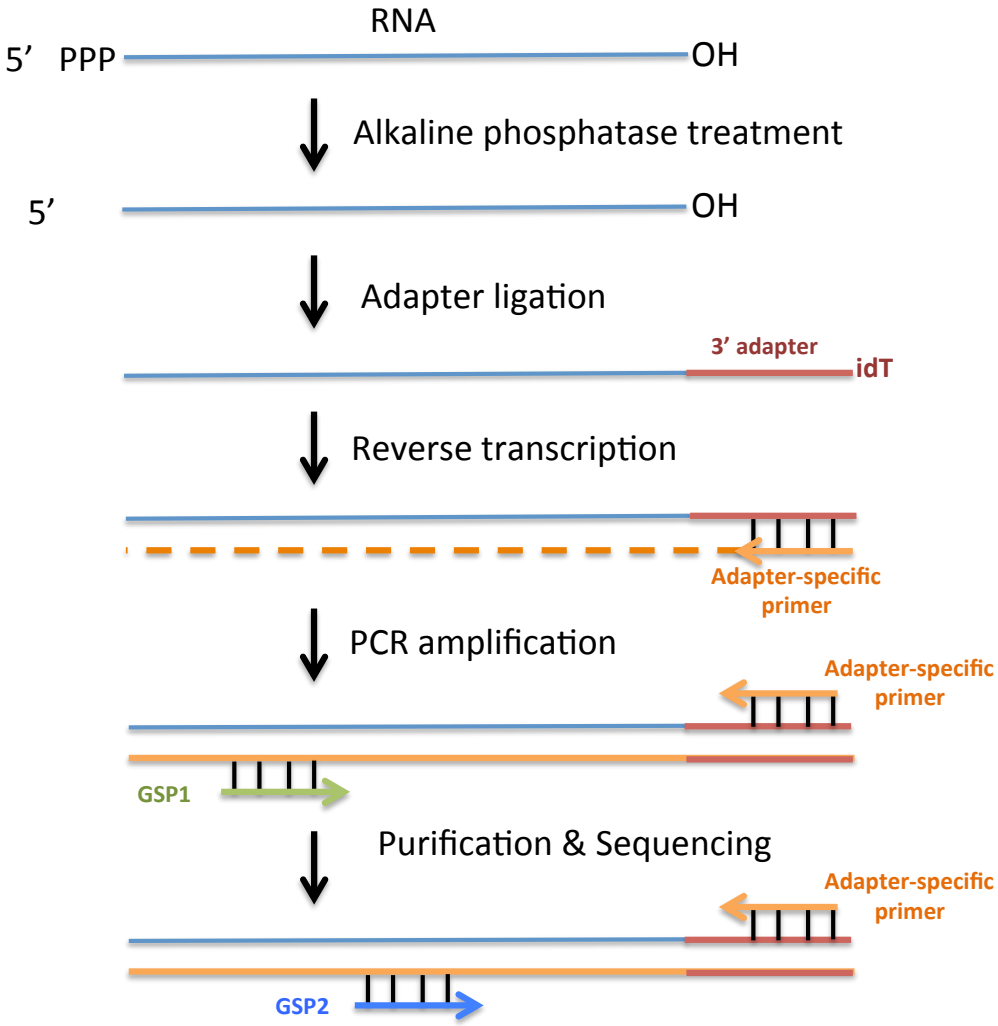
** The fold change is calculated as the ratio of the intensities between the treated condition and the untreated control. Thus a fold change > 1 indicates that the sRNA is significantly more expressed in the treated condition.

****Coordinates validated by 5´and 3´-RACE experiments.

Supplementary Figure 1. 5´RACE strategy used to identify 5´ends of primary and processed transcripts.



Supplementary Figure 2. 3´RACE strategy used to identify 3´ends of transcripts.



Article 4

Small RNA OsiS links oxidative stress to quorum sensing control in *Pseudomonas aeruginosa*

Gómez-Lozano, M. & Molin, S.

Manuscript in preparation, **2013**.

Small RNA OsiS links oxidative stress to quorum sensing control in *Pseudomonas aeruginosa*

María Gómez-Lozano[1,2] and Søren Molin[1,2]*

[1] *Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.*

[2] *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.*

*To whom correspondence should be addressed (e-mail: sm@bio.dtu.dk).

ABSTRACT

Small RNAs are known to control diverse adaptation processes, where bacterial physiology is adjusted in response to environmental cues. In this study we investigated the role of sRNA OsiS that was recently identified in genome-wide search of sRNAs in *P. aeruginosa*. OsiS is highly transcribed during oxidative stress conditions. We show that by inducing the expression of OsiS the levels of the sRNA PhrS are greatly reduced. PhrS activates the translation of PqsR under low oxygen concentrations, which in turn activates the synthesis of PQS. Thus, OsiS links the oxygen levels to the production of quorum sensing (QS) molecules. We hypothesize that the interaction is by direct base-pairing between the two sRNAs, with a predicted recognition site of OsiS at the highly conserved-region of PhrS. However, more experiments are required to know the exact nature of the interaction between these two sRNAs. OsiS is, to the best of our knowledge, the first sRNA whose main function seems to be regulating the cellular levels of another sRNA.

INTRODUCTION

The opportunistic human pathogen *Pseudomonas aeruginosa* can cause severe infections, particularly in immuno-compromised and cystic fibrosis (CF) patients. Despite aggressive antibiotic therapy, a high number of CF patients suffer from chronic *P. aeruginosa* infections. The establishment of a chronic infection requires a high level of adaptation of the pathogen to the host environment (1,2).

During the last decade many novel small regulatory RNAs (sRNAs) have been described

in different bacteria, including *Escherichia coli, Listeria monocytogenes, Staphylococcus aureus, Helicobacter pylori, Vibrio cholerae* and *Pseudomonas aeruginosa* (3-8). In *P. aeruginosa*, sRNAs are known to be involved in the regulation of virulence genes (RsmY/Z), carbon catabolite repression (CrcZ), iron metabolism (PrrF1, PrrF2, PrrH), and quorum sensing (PhrS) (9-13). Recently, we identified more than 500 novel intergenic sRNAs (8) and more than 200 novel antisense transcripts (unpublished data) using different RNA-Seq library protocols (14). Following this study, we used RNA-Seq to study the transcriptional response of *P. aeruginosa* under different stress conditions, including osmotic, oxidative and antibiotic stress (unpublished data). In this study we show that one of the novel sRNAs (named pant67 in (8) and OsiS in this study) is highly expressed during oxidative stress, and that its role is to downregulate the sRNA PhrS.

PhrS is a 213-nt long sRNA, predicted by biocomputation (15,16) and isolated by an RNomics approach (7). PhrS is highly transcribed under low oxygen conditions, and its expression requires the oxygen-responsive regulator ANR (13). PhrS activates the translation of the *pqsR* gene by base-pairing with a short untranslated open reading frame (uof) to which the *pqsR* gene is translationally coupled (13). PqsR is one of the key QS regulators in *P. aeruginosa* and its synthesis induces the PQS synthetic operon *pqsABCDE* (17,18). Beside its function in inter-bacterial signaling, PQS is required for full synthesis of the *P. aeruginosa* virulence factor pyocianin, which can react with molecular oxygen to generate oxidative stress (19). In addition to its function as a regulatory RNA, PhrS encodes a highly conserved 37 amino-acid peptide (7). The peptide contains a predicted transmembrane segment and sub-cellular fractionation revealed that the peptide is indeed located in the cytoplasmic membrane. Sonnleitner *et al.* (2011) performed a BLASTN search with the DNA sequence of PhrS to identify homologues. This analysis revealed homologues of PhrS only in 12 different isolates of *P. aeruginosa* (13). Sonnleitner *et al.* (2011) performed another BLAST search with the amino acid sequence of the peptide encoded by PhrS, and found homologous amino acid sequences in several Pseudomonas species as well as in *A. vinelandii* (13). This analysis identified a highly conserved region of 12 nt downstream of the internal ORF of PhrS (13). In this study we show that the transcription of OsiS RNA is activated during oxidative stress, and that by inducing the expression of OsiS the levels of the sRNA PhrS are greatly reduced.

MATERIALS AND METHODS

Primers and adapters

Supplementary Table S1 lists all primers and adapters used in this study.

Plasmid construction

The plasmids used in this study are listed in Table 1. Plasmids for L-arabinose-inducible expression of asRNAs were constructed by amplification of plasmid pBAD33 (20) with primers F1-SacI and R1 producing a PCR product of 1250 bp (F1-SacI introduces a SacI restriction site). For amplification of the asRNA insert, genomic DNA from strain

PAO1 was amplified with primers F2 and R2-XhoI (F2 starts with the sRNA +1 site and R2-XhoI introduces an XhoI restriction site). The 5´ends of primers F2 and R1 are complementary to each other. PCR synthesis by overlap extension (PCR-SOE) was carried out by combining the 2 previously amplified products and running PCR for 10 cycles, after which primers F1-SacI and R2-XhoI are added and the reaction continues for another 30 cycles. The obtained product was digested with SacI and XhoI and cloned into the MCS of plasmid mini-CTX2 (21).

Bacterial strains

The bacterial strains used in this study are listed in Table 1. The *Pseudomonas aeruginosa* PAO1 strain was used as wild-type strain. The *E. coli* CC118*λpir* strain was used as cloning host (22). Strains pBAD-*osiS* (Figure 2A) and pBAD-control were created by integration of plasmids pBAD-*osiS* and pBAD-control, respectively, at the *attB* site of strain PAO1, followed by Flp-mediated excision of the plasmid backbone, as described in (21). The first nucleotide of the transcribed product is the +1 of the primary OsiS. This is an important consideration when cloning noncoding RNAs under the control of heterologous promoters, since altering the natural RNA start site may have unintended consequences for its function. The *attB* region was sequenced using primers P_{ser}-up and P_{ser}-down to make sure that the integration was successful and the plasmid backbone excised.

Table 1. Strains and plasmids used in this study.

Strain or plasmid	Relevant markers/genotype	Source or reference
Strains		
P. aeruginosa PAO1	Wild type	Laboratory strain
P. aeruginosa pBAD-*osiS*	pBAD-*osiS* fragment integrated at the *attP* site	This study
P. aeruginosa pBAD-control	pBAD-control fragment integrated at the *attP* site	This study
E. coli CC118*λpir*	*araD139 Δ(ara leu)7697 ΔlacX74 phoA20 galE galK thi rpsE rpoB argE(Am) recA1 λpir* lysogen	Herrero *et al* (1990)
Plasmids		
mini-CTX2	TcR; self-proficient integration vector with *tet*, Ω-FRT-attP- MCS, *ori*, *int*, *oriT* and *lacI*	Hoang *et al* (2000)
pBAD33	CmR, pBAD expression plasmid with pBR322	Guzman *et al* (1995)
pBAD-*osiS*	TcR; OsiS expression plasmid, *osiS* is controlled by the plasmid-borne PBAD promoter	This study
pBAD-control	TcR; pBAD control plasmid, expresses 322 nt nonsense RNA followed by a terminator	This study
pFLP2	AmpR, source of Flp recombinase	Hoang *et al* (1998)

Growth conditions

Growth in Luria–Bertani (LB) broth (250 rpm, 37°C) or on LB plates at 37°C was used throughout this study.

Stress exposure experiments. Overnight cultures of *Pseudomonas aeruginosa* PAO1 were diluted to a starting OD_{600} of 0.01 and grown to an OD_{600} of 0.5, at which 3xMIC ciprofloxacin or 1mM hydrogen peroxide was added. Concentrated stock solutions of H_2O_2 and ciprofloxacin were prepared fresh daily. Cells were harvested 1 hour after the addition of ciprofloxacin or hydrogen peroxide. RNA was extracted and used to make RNA-Seq libraries LIB<500. Experiments were performed in duplicates.

Pulse-expression experiments. Overnight cultures of strains pBAD-*osiS* were diluted to a starting OD_{600} of 0.01 and grown to an OD_{600} of 0.5, at which sRNA expression was induced with 1% L-arabinose. Cells were harvested after 30 min and 45 min of the induction. RNA was extracted and used to make RNA-Seq libraries LIB>100. Two sets of negative-control conditions were used for comparison. In one, the control strain pBAD-control was induced with arabinose; in the second, cells containing the pBAD-*osiS* construct were grown without arabinose induction. Experiments were performed in duplicates.

RNA isolation

Harvested cells were mixed immediately with 0.2 volumes of STOP solution (95% ethanol, 5% phenol) and pelleted by centrifugation. Total RNA was extracted with Trizol (Invitrogen). Removal of DNA was carried out by treatment with DNase I (Fermentas) in combination with the RNase inhibitor RiboLock (Fermentas). The integrity of total RNA and DNA contamination were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

Removal of 23S, 16S and 5S rRNAs

The 23S, 16S and 5S rRNAs were removed by subtractive hybridization using the MICROBExpress kit (Ambion) with modifications as previously described (8) (14). Capture oligonucleotides complementary to the rRNAs of *P.aeruginosa* PAO1 were used (Supplementary Table S1). Briefly, 5 µM of each capture oligonucleotide was used, for a total capture oligonucleotide concentration of 25 µM.

Compared with the standard protocol, 25% more capture oligonucleotides and magnetic beads were used. Removal of rRNAs was confirmed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

Library preparation and RNA sequencing

Two different sequencing libraries (LIB>100 and LIB<500) were constructed as previously described (8,14). Each library type was prepared with RNA isolated from cells grown in duplicate for each condition studied. After each step the samples were validated using an Agilent 2100 Bioanalyzer (Agilent Technologies), and the final concentration was measured using a Qubit 2.0 Fluorometer (Invitrogen). The libraries were sequenced using the Illumina HiSeq2000 platform with a paired-end protocol and read lengths of 100 nt.

LIB>100 library. This type of library contains transcripts longer than 100 nt and was prepared using the TruSeq RNA Sample Preparation kit (Illumina) as previously described (8,14). Briefly, the rRNA-depleted RNA is fragmented using divalent cations under elevated temperature. The cleaved RNA fragments are copied into cDNA using reverse transcriptase and random primers, followed by second-strand cDNA synthesis using DNA polymerase I and RNase H. After this step, transcripts shorter than 100 nt are not retained due to the purification performed with the Agencourt AMPure XP beads (Beckman Coulter Genomics). The cDNA fragments then go through an end repair process, 3'-addition of single 'A' bases and adapter ligation. This is followed by product purification and PCR amplification to generate the final cDNA library.

LIB<500 library. This type of library is strand-specific and contains transcripts shorter than 500 nt. RNA size selection was performed by running total RNA on 10% polyacrylamide gels containing 10 M urea. Gel slices corresponding to RNAs up to 500 nt were excised, followed by elution of RNA in 0.4 M NaCl and precipitation with ethanol. The 5S rRNA was depleted as previously described, followed by treatment with Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37°C for 90 min. Tobacco Acid Pyrophosphatase (TAP) is used to convert 5′-triphosphate RNA into 5′-monophosphate RNA, which is important for correct adapter ligation. This was followed by treatment with RNase III (Ambion) for 10 min at 37°C to fragment the RNA. RNase III fragments RNA into smaller pieces containing a 5′-phosphoryl group and a 3′-hydroxyl group, which is important for specific adapter ligation in the next step. Sequential ligation of RNA 3′ and 5′ adapters was performed using the adapters and enzymes from the TruSeq Small RNA Sample Preparation kit (Illumina). Next, reverse transcription followed by PCR amplification was performed to form cDNA constructs based on the RNA fragments ligated with 3′ and 5′ adapters, selectively enriching fragments with adapter molecules on both ends. The reverse transcription and subsequent PCR amplification were performed using the enzymes and reagents from the TruSeq Small RNA Sample Preparation kit (Illumina). Agencourt AMPure XP beads (Beckman Coulter Genomics) were used for the post-PCR clean-up. In this type of library the fragments corresponding to transcripts shorter than 100 nt are retained due to ligated RNA adapters that increase the length of the fragments by 125 nt.

Data analysis

Our analysis pipeline is described in detail in Gómez-Lozano et al (2013) (14).

Briefly, sequencing reads were mapped onto the *P. aeruginosa* PAO1 genome using the Bowtie 2 short read aligner (23). Read alignments from Bowtie 2 were handled using SAMtools (24). In order to obtain normalized expression intensities of the read coverage depth at each position in the genome, the number of reads in each replicate was normalized according to the total number of reads in the library, and expression intensities from replicate samples were averaged. A two-sample Student's *t*-test was performed on the average expression of the transcripts to determine those with differential expression between the conditions tested (*P*-value < 0.05).

Rapid Amplification of cDNA Ends (RACE)

Schematic representations of the 5´ and 3´ RACE protocols can be seen in Supplementary Figures 1 and 2. Supplementary Table 1 lists primers and adapters used.

5´ RACE. 2 µg rRNA-depleted RNA was incubated with 10 U of Tobacco Acid Pyrophosphatase (Epicentre Technologies) at 37ºC for 1 h to convert RNA 5´ triphosphates in 5´ monophosphates. The same amount of control RNA was incubated under the same conditions in the absence of the enzyme. Reactions were stopped by phenol-chloroform extraction, followed by etanol-sodium acetate precipitation. Precipitated RNAs were redissolved in water, mixed with 500 pmol of 5´ RNA adapter, heat-denatured at 95ºC for 5 min, then quick-chilled on ice. A short RNA adapter was ligated was ligated with 50 U of T4 RNA ligase (Thermo Scientific) at 37ºC for 1 h. Reactions were stopped by phenol-chloroform extraction, followed by ethanol-sodium acetate precipitation. Precipitated RNAs were re-dissolved in 20 µl water. Then 10 µl of 5´adapter-ligated RNA was reverse-transcribed using 2 pmol of primer complementary to the sRNA (5´-GSP1) and the Thermoscript RT-PCR system (Invitrogen) according to the manufacturer's instructions. Reverse transcription (RT) was performed in three

subsequent 20 min steps at 55ºC, 60ºC, and 65ºC, followed by treatment with RNase H. Primers were removed using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The products of RT were amplified using 10 pmol of another primer complementary to the sRNA (5´-GSP2) and 10 pmol the 5´ adapter-specific primer, together with the Maxima Hot Start PCR Master Mix (Thermo Scientific) according to the manufacturer's instructions. Negative controls were performed using the 5´ adapter-ligated RNA as template. The PCR products were resolved and purified using E-Gel SizeSelect 2% Agarose gels (Invitrogen). Products were sequenced with 5´-GSP2 and 5´ adapter-specific primers by LGC Genomics GmbH (Germany).

3´ RACE. 2 µg rRNA-depleted RNA was dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) at 37ºC for 1 h. Reactions were stopped by phenol-chloroform extraction, followed by etanol-sodium acetate precipitation. Ligation was done as above with a 3´ RNA adapter with a 3´-inverted deoxythymidine (3´-idT). RT was carried out as described, but with 10 pmol of a single primer complementary to the 3´ RNA adapter. PCR amplification, band purification and sequence analysis was done as described above. All enzymatic treatments of RNA were performed in the presence of 20 units of RiboLock RNase Inhibitor (Thermo Scientific).

Quantitative reverse transcriptase PCR (qRT-PCR)

DNase-treated total RNA was reverse transcribed using gene-specific primers and the ThermoScript RT-PCR System (Invitrogen). Quantitative PCR was run using the Brilliant III Ultra-Fast SYBR Green qPCR Master Mix (Stratagene), the Strategene Mv3005P equipment and MxPro qPCR software. 5S rRNA gene was used as the normalizer. Three independent sets of experiments were performed. In all cases, controls lacking reverse transcriptase were included to assess DNA contamination.

RESULTS

OsiS is expressed under oxidative stress conditions.

In a recent study, we found more than 500 novel intergenic sRNAs (8) and more than 200 novel antisense transcripts (unpublished data) using different RNA-Seq library protocols (14). Following this study, we used RNA-Seq to study the transcriptional response of *P. aeruginosa* under different stress conditions, including osmotic, oxidative and antibiotic stress (unpublished data). It was discovered that one of the novel sRNAs, OsiS, is highly expressed following ciprofloxacin (CIP) and hydrogen peroxide (H_2O_2) treatment. Ciprofloxacin is a fluoroquinolone antibiotic commonly used against *P. aeruginosa* infections. Ciprofloxacin inhibits DNA topoisomerase II and DNA topoisomerase IV activities, eventually leading to bacterial cell death. In addition, an increase of reactive oxygen species (ROS) in the bacterial cells in response to ciprofloxacin has been shown (25,26). Hydrogen peroxide is a common ROS produced by the host immune system, which *P. aeruginosa* exogenously encounters during infection (27-29). In our experiments, OsiS expression was not activated during exposure to NaCl, ß-lactams (piperacillin, ceftazidime, aztreonam, meropenem), aminoglycosides (tobramycin), macrolides (azithromycin), colistin or tetracycline. These results indicate that OsiS is activated during oxidative stress, as caused by H_2O_2 and ciprofloxacin. As seen in Figure 1B the levels of OsiS are higher after CIP exposure (fold change 99.99±36.37) than after H_2O_2 (fold change 9.34±0.04). This is most likely due to the concentration of CIP and H_2O_2 used in the experiments, which was lethal in the case of CIP (3xMIC, 0,375 µg/ml) and sub-lethal in the case of H_2O_2 (1 mM).

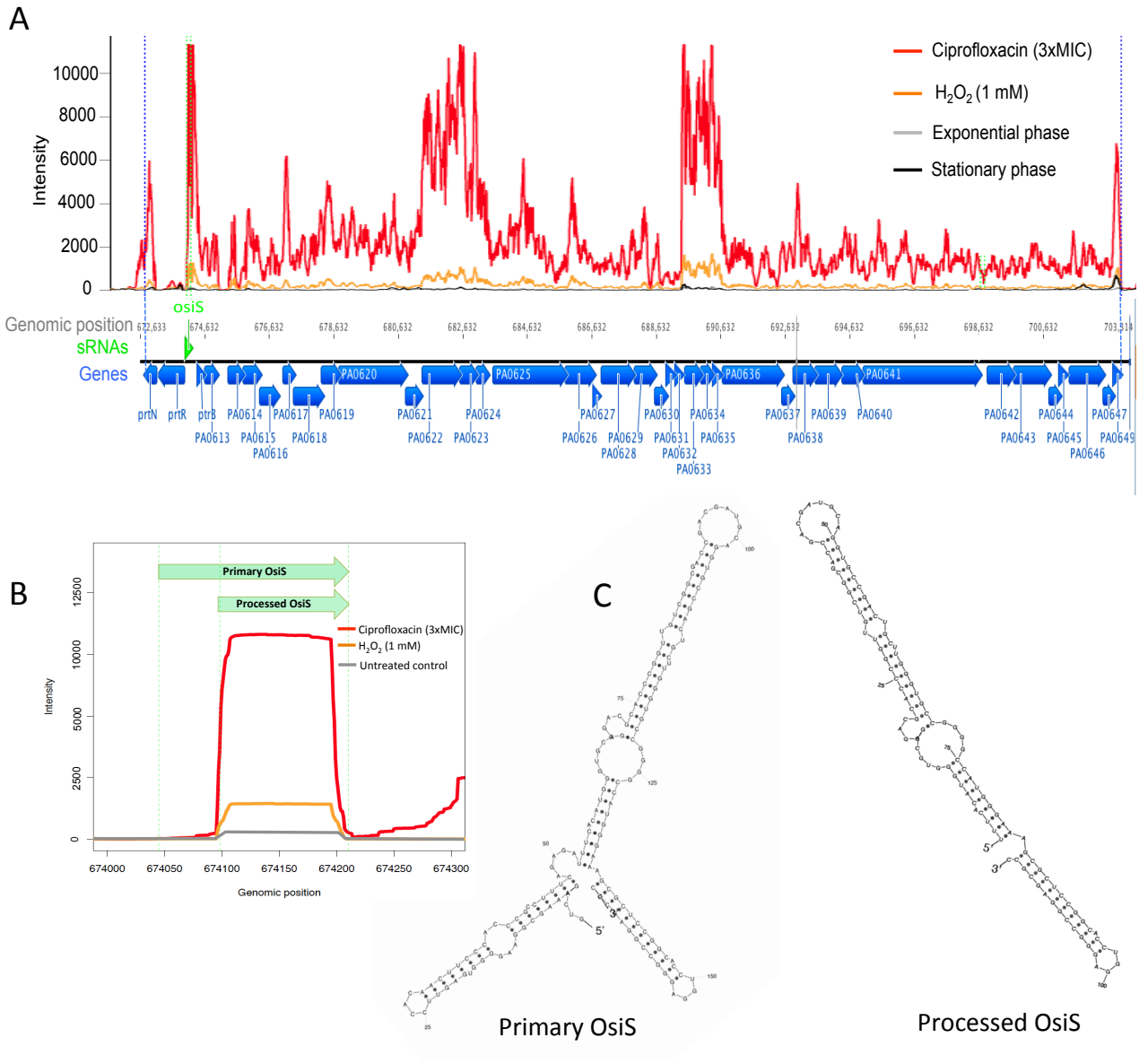


Figure 1. Changes in transcriptional response upon ciprofloxacin (3xMIC) and hydrogen peroxide (1 mM) treatment. A. The PAO1 R- and F- type pyocin loci. All the pyocin locus, with the exception of the pyocin repressor gene *prtR*, are upregulated upon ciprofloxacin and hydrogen peroxide treatment. The *osiS* gene (green arrow) is located between the genes *prtR* and *prtB*. **B.** Changes in expression of OsiS upon ciprofloxacin (3xMIC) and hydrogen peroxide (1 mM) treatment. The coordinates of the primary and processed versions of OsiS (assessed by 5´-RACE) are marked with green arrows. **C.** Predicted secondary structures of OsiS (primary and processed) predicted with sfold (30).

Pulse-expression of OsiS leads to down-regulation of PhrS.

The *osiS* gene is located in between the genes *prtR* and *prtB* in the R- and F- type pyocin loci. The R- and F- type pyocin loci are upregulated upon H_2O_2 and CIP treatment with fold changes similar as those of OsiS (Figure 1A). Pyocins are bacteriocins produced by *P. aeruginosa* that have been shown to kill both a diversity of *P. aeruginosa* strains in addition to a number of Campylobacter species *N. gonorrhea, N.*

meningiditis, H. ducreyi, P. fluorescens, P.putida and *B. cepacia* complex (31-36). Some *P. aeruginosa* strains do not produce both R- and F- types of pyocins, however all known strains produce at least one of the two types (37,38). DNA damage, and therefore the bacterial SOS response, is believed to trigger the induction of pyocins. Some studies have shown that the R- and F- type pyocin genes are upregulated during both lethal and sublethal treatment with CIP (39) and during H_2O_2 exposure (40).

Based on the above mentioned, our initial hypothesis was that OsiS was involved in the regulation of pyocin production during oxidative stress conditions. To investigate this further, OsiS was pulse-expressed with an inducible promoter and its immediate effects studied by RNA-Seq. In these experiments OsiS expression was restricted to a construct inserted at the *attB* site containing *osiS* under the control of the heterologous pBAD promoter, which is only induced in the presence of arabinose (Figure 2A). Inserting the construct at a specific site of the genome has several advantages over using a plasmid-based system. Since the integration vector used in this study is equipped with FRT sites (see Materials and Methods), it allows Flp-mediated excision of the plasmid backbone. Thus, our engineered strains are devoid of any selection marker and the experiments can be done without using any selective pressure. Different copy-numbers in the strains to be compared can be an issue when working with plasmid-based system, but that represents no problem when working with chromosomal integrations. Furthermore, genomic DNA and plasmid DNA differ in their degree of supercoiling, a state of DNA that has been shown to play a major role in regulation of gene expression (41).

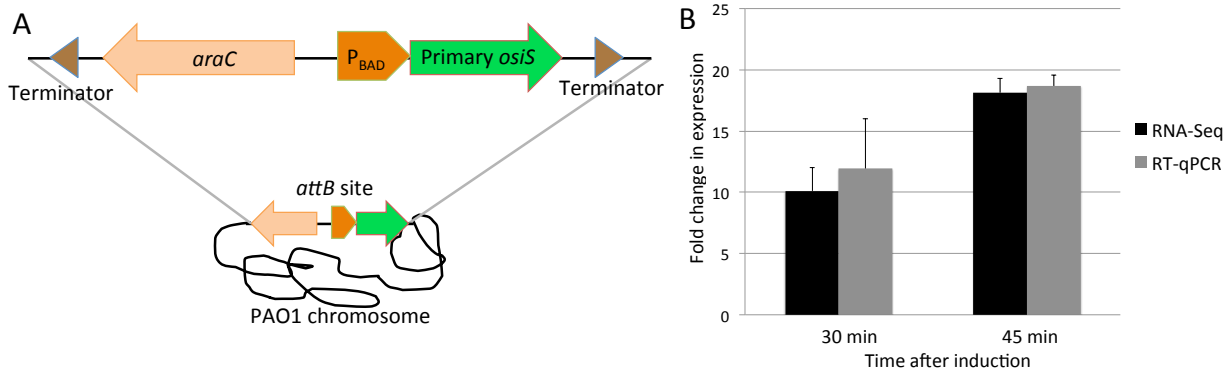


Figure 2. A. The pBAD-*osiS* construct inserted at the *attB* site of strain pBAD-*osiS*. The first nucleotide of the transcribed product is the +1 of the primary OsiS. **B.** Fold change in OsiS levels after induction with L-arabinose. The levels of OsiS were measured both with RNA-Seq and RT-qPCR. 1% L-arabinose was added to induce the expression of OsiS when OD_{600} reached 0.5. The OsiS level was detected after the expression of OsiS was induced for 30 and 45 min. Strain pBAD-control was used as the control. For the RT-qPCR experiments, the expression was normalized to the level of 5S rRNA in each sample.

In pulse-expression analysis, the first nucleotide of the transcribed product should be the +1 of the primary transcript to be pulse-expressed. This is an important consideration when cloning sRNAs under the control of heterologous promoters, since altering the natural RNA start site may have unintended consequences for its function. Thus, 5´- and 3´-RACE experiments were performed to accurately map the boundaries of OsiS. From the RNA-Seq data OsiS appeared as a 113-nt transcript (674,098-674,210) (Figure 1B). However, differential 5´-RACE experiments revealed that the start site of OsiS was located 53 bp upstream. This result fitted well with the prediction of -10 and -35 σ^{70} promoter elements (Supplementary figure S3). 3´-RACE results were consistent with RNA-Seq data. Thus, the primary transcript of OsiS is 166-nt long (674,045-674,210). The predicted structures of both primary and processed OsiS are shown in Figure 1C. The RNA-Seq data revealed that OsiS was already processed after 30 min of pulse-expression of the primary transcript of OsiS.

Cells carrying the pBAD-*osiS* construct were grown to mid-log phase in LB and treated with 1% arabinose for 30 and 45 min, respectively. Two sets of negative-control conditions were used for comparison. In one, the control construct, pBAD-control was induced with arabinose; in the second, cells containing the pBAD-*osiS* construct were grown without arabinose induction. Total RNA was extracted from all cultures and processed for RNA-Seq analysis (see Materials and Methods). The gene expression profiles were compared in two ways: (i) induced pBAD-*osiS*-containing cells were compared to induced pBAD-control-containing cells, and (ii) uninduced pBAD-*osiS*-containing cells were compared to induced pBAD-osiS-containing cells. The results for the two types of comparisons showed consistent results, with the exception of the expected differences for arabinose-inducible genes. The expression levels of OsiS obtained after induction were investigated both by RT-PCR and RNA-Seq (Figure 2B), obtaining similar results. The fold change of OsiS 30 and 45 min after the induction were 10.2 and 18.0, respectively. These fold changes are slightly higher than those obtained upon treatment with 1 mM H_2O_2.

Table 2 shows the list of genes significantly up- or downregulated after 30 and 45 min of pulse-expression of OsiS. The main effect after 30 min was the dramatic downregulation of PhrS levels (fold change -79.7). PhrS activates the translation of the *pqsR* gene. PqsR is one of the key QS regulators in *P. aeruginosa* and its synthesis induces the PQS synthetic operon *pqsABCDE* (17,18). In fact, the *pqsABCDE* genes are downregulated 45 min after OsiS induction (fold changes between -1.6 and -2.1). Some of the genes controlled by PqsR are also downregulated upon OsiS pulse-expression, with greater effects after 45 min than after 30 min: the phenazines biosynthesis genes *phzA1B1*, required for pyocianin production (42), and the *phnAB* genes, encoding the enzymes required to convert chorismate to anthranilate, the PQS precursor (43). Some of

the genes required for exoenzyme S production are also dowregulated. Interestingly, the genes *rhlI* and *rhlR* are also mildly downregulated in our experiments, and it has been shown that PQS induces the transcription of genes *rhlI* and *rhlR* (44). The outer membrane protein OprG was also more than 2-fold downregulated. Strikingly, only one gene, the polyhydroxyalkanoate synthesis gene *phaF*, increased its levels after OsiS overexpression. The levels of sRNA P5 are also reduced significantly, but unfortunately the role of this sRNA is still unknown. These results indicate that OsiS provides a link between oxygen levels and QS, by downregulating the levels of the sRNA PhrS during oxidative stress conditions.

Table 2. Transcripts up- and downregulated upon *osiS* overexpression.

Locus ID	Gene	Description	Pathway	Fold change after 30 min	Fold change after 45 min
Transcripts upregulated upon *osiS* overexpression					
PA0611.1	*osiS*	OsiS	Non-coding RNA gene	+10.2	+18
PA5060	*phaF*	polyhydroxyalkanoate synthesis protein	Central intermediary metabolism	+1.6	+1.9
Transcripts downregulated upon *osiS* overexpression					
PA1003	*mvfR/pqsR*	Transcriptional regulator MvfR	Transcriptional regulators	-1.3	-1.4
PA3476	*rhlI*	autoinducer synthesis protein RhlI	Adaptation, Protection	-1.5	-1.4
PA1713	*exsA*	transcriptional regulator ExsA	Protein secretion/export apparatus; Transcriptional regulators	-1.6	-1.5
PA1002	*phnB*	anthranilate synthase component II	Amino acid biosynthesis and metabolism; Adaptation, Protection	-1.2	-1.5
PA3477	*rhlR*	transcriptional regulator RhlR	Adaptation, Protection; Transcriptional regulators	-1.3	-1.5
PA1001	*phnA*	anthranilate synthase component I	Adaptation, Protection	-1.2	-1.5
PA1712	*exsB*	exoenzyme S synthesis protein B	Protein secretion/export apparatus	-1.4	-1.6
PA1711	*exsE*	ExsE	Transcriptional regulators; Protein secretion/export apparatus	-1.6	-1.6
PA1000	*pqsE*	Quinolone signal response protein	Biosynthesis of cofactors, prosthetic groups and carriers	-1.2	-1.6
PA0999	*pqsD*	3-oxoacyl-[acyl-carrier-protein] synthase III	Biosynthesis of cofactors, prosthetic groups and carriers	-1.1	-1.7
PA0998	*pqsC*	PqsC	Biosynthesis of cofactors, prosthetic groups and carriers	-1.1	-1.7
PA0996	*pqsA*	probable coenzyme A ligase	Biosynthesis of cofactors, prosthetic groups and carriers	1	-1.7
PA1714	*exsD*	ExsD	Transcriptional regulators; Protein secretion/export apparatus	-1.5	-2.1
PA0997	*pqsB*	PqsB	Biosynthesis of cofactors, prosthetic groups and carriers	1	-2.1
PA1433	-	conserved hypothetical protein	Membrane proteins	-1.2	-2.3
PA4132	-	conserved hypothetical protein	Hypothetical, unclassified, unknown	-1.8	-2.3
PA4067	*oprG*	Outer membrane protein OprG precursor	Membrane proteins	-2.3	-2.3
PA1699	*pcr1*	Pcr1	Hypothetical, unclassified, unknown	-1.4	-2.4
PA5480	-	hypothetical protein	Hypothetical, unclassified, unknown	-3.7	-2.4
PA4134	-	hypothetical protein	Hypothetical, unclassified, unknown	-1.7	-2.4
PA4211	*phzB1*	probable phenazine biosynthesis protein	Secreted Factors (toxins, enzymes, alginate)	-1.3	-2.5
PA0026	*plcB*	phospholipase C, PlcB	Hypothetical, unclassified, unknown	-1.3	-2.8
PA4210	*phzA1*	probable phenazine biosynthesis protein	Secreted Factors (toxins, enzymes, alginate)	-2.2	-2.8
PA2759	-	hypothetical protein	Hypothetical, unclassified, unknown	-6.3	-3
PA2501	-	hypothetical protein	Membrane proteins	-7.4	-3.6
PA0836.1	*P5*	P5	Non-coding RNA gene	-2.8	-6.6
PA1673	-	hypothetical protein	Hypothetical, unclassified, unknown	-11.4	-11.1
PA0200	-	hypothetical protein	Hypothetical, unclassified, unknown	-32.3	-18.1
PA1414	-	hypothetical protein	Hypothetical, unclassified, unknown	-22.6	-19
PA3305.1	*phrS*	PhrS	Non-coding RNA gene	-79.7	-38.9

OsiS conservation and interaction with PhrS

It is clear from the transcriptomic results that OsiS lowers the levels of PhrS. However, the question of how these two sRNAs interact is not answered by the RNA-Seq data. The interaction could be by direct base-pairing between the sRNAs, or OsiS could be directly affecting some other transcript or protein that regulates PhrS. Sonnleitner *et al.* (2011) performed a BLASTN search to identify PhrS homologues. This analysis revealed homologues of PhrS in 12 different isolates of

P. aeruginosa (13). Interestingly, the same kind of analysis indicated that OsiS is conserved in the exact same 12 isolates (Figure 3). This indicates that OsiS and PhrS have evolved together in the same isolates.

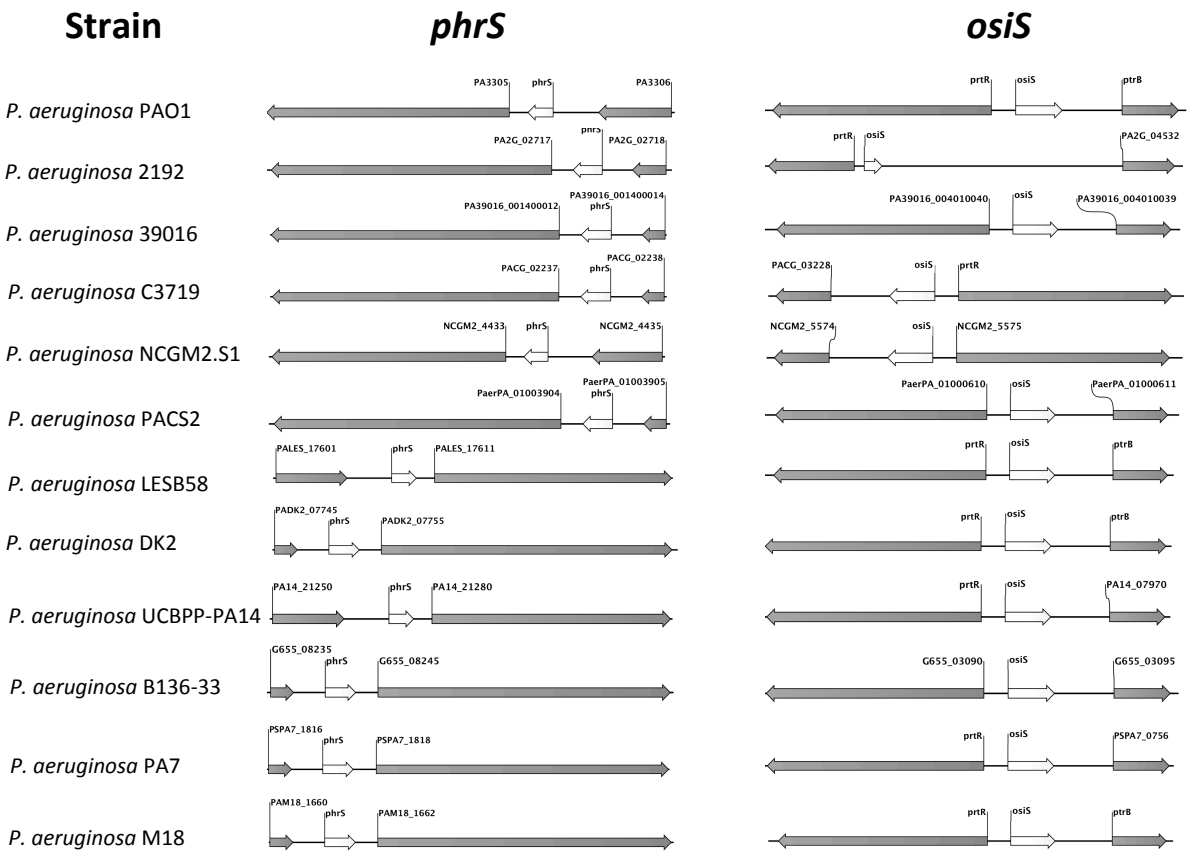


Figure 3. Genetic organization of *phrS* and *osiS* regions in all the *P. aeruginosa* strains in which these sRNAs are found. White arrows represent the sRNAs and grey ones represent flanking genes.

To identify a potential annealing site between OsiS and PhrS the IntaRNA program (45) was used. As shown in Figure 4, this analysis predicted a recognition site of OsiS upstream and at the highly conserved-region of PhrS. This highly conserved region of PhrS also base-pairs with the leader sequence of the upstream open reading frame (uof), which is co-transcribed with *pqsR*, activating the translation of PqsR (13). This suggests that the interaction between OsiS and PhrS might be by direct base-pairing. However, more experiments are required to know the exact nature of the interaction between these two sRNAs.

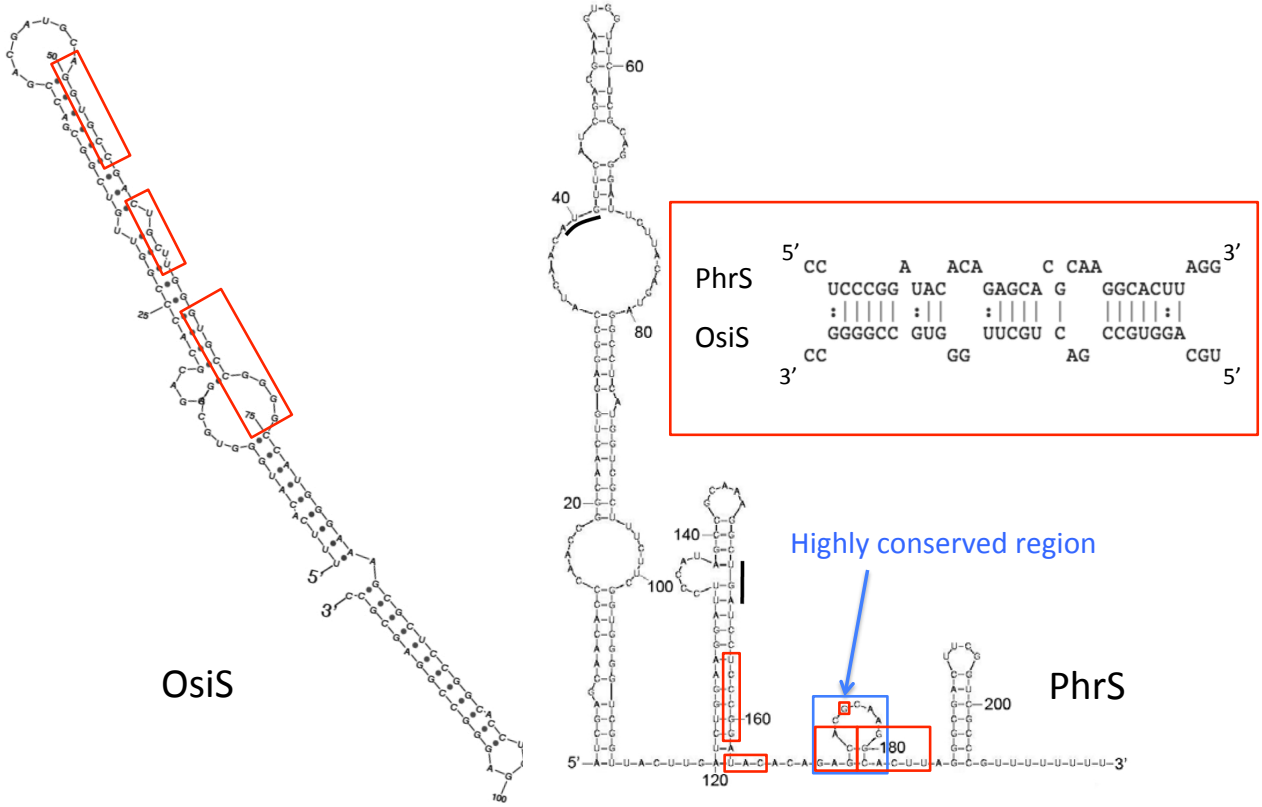


Figure 4. Predicted interaction sites between OsiS and PhrS. The secondary structures of OsiS and PhrS are shown on the left. The structure of PhrS is adapted from (13). The predicted interaction sites were predicted by the IntaRNA algorithm (45) and are boxed in red. The base-pairing interaction is depicted on the right. Part of the highly conserved region of PhrS is required for the interaction. The start and stop codons of the 37-aa peptide encoded by PhrS are underlined in black.

DISCUSSION

Under conditions that generate oxidative stress, OsiS is highly expressed. We have shown that 30 min of induction of the expression of OsiS produces a significant downregulation in the levels of PhrS. Also, many of the downregulated genes after 45 min of induction of OsiS are PqsR-dependent. These results indicate that OsiS provides a link between oxygen levels and QS, by downregulating the levels of the sRNA PhrS during oxidative stress conditions. Thus, OsiS is another sRNA that responds to oxygen levels in bacteria, such as PhrS in *P. aeruginosa* (13), RgsA in *P. fluorescens* (15), OxyS, FnrS and MicF in *E. coli* (46-49), and RliB in *L. monocytogenes* (50).

The *osiS* gene is located in the R- and F- type pyocin loci. DNA damage, and therefore the bacterial SOS response, triggers the induction of pyocins. Some studies have shown that the R- and F- type pyocin genes are upregulated during both lethal and sublethal treatment with CIP (39) and during H_2O_2 exposure (40). In our study the R- and F- type pyocin loci are upregulated upon H_2O_2 and CIP treatment with

fold changes similar as those of OsiS. However, OsiS does not seem to be involved in the regulation of pyocin production. It is interesting to note that OsiS is located in a genomic region that gets transcriptionally active during the same conditions as OsiS.

Figure 5 shows our current model of the regulatory function of OsiS. During oxidative stress conditions the sRNA OsiS is transcribed at high levels. The high levels of OsiS induce the downregulation of the sRNA PhrS. In addition, the oxygen-responsive regulator ANR, which is required for PhrS transcription, is not active. On the contrary, during low oxygen conditions, OsiS is not transcribed, ANR is active and PhrS is highly transcribed (7,13). PhrS activates the translation of the *pqsR* gene. PqsR is one of the key QS regulators in *P. aeruginosa* and its synthesis induces the PQS synthetic operon (17,18). The PhrS-mediated stimulation of PQS synthesis leads to an increased production of pyocianin. Pyocianin is a virulence factor that can be reduced by NAD(P)H or glutathione, and in its reduced form can react with oxygen, forming ROS (19). The ROS produced by the synthesis of pyocianin might be a signal that activates OsiS transcription and ANR inactivation, in order to regulate the levels of sRNA PhrS and in consequence the production of PQS.

We have shown that expression of OsiS results in downregulation of PhrS levels, but how do these two sRNAs interact with each other? The interaction could be by direct base-pairing, or OsiS could be directly affecting some other transcript or protein that, in turn, downregulates PhrS. The DNA sequences of OsiS and PhrS are conserved in the same strains of *P. aeruginosa*, indicating that these two sRNAs might have evolved in parallel. We have predicted a recognition site of OsiS upstream and at the highly conserved-region of PhrS. This highly conserved region of PhrS has been shown to base-pair with the leader sequence of the upstream open reading frame of *pqsR*, activating the translation of PqsR (13). Based on this, we hypothesize that OsiS and PhrS interact by direct base-pairing. However, more experiments are required to know the exact nature of the interaction between these two sRNAs. 5´-RACE analysis has shown that OsiS is found in the cell as a processed transcript, which lacks a hairpin of 53 nt at the 5´ end with respect to the primary transcript of OsiS. It would also be interesting to know whether the primary or the processed version of OsiS is the one responsible for the PhrS downregulation. The oxygen-responsive regulator ANR controls the synthesis of PhrS by promoting its transcription during low-oxygen conditions, but is the transcription of PhrS also affected by OsiS in the presence of high-oxygen tension? Or does OsiS affect the stability of PhrS? Also, is the Hfq chaperone required for the interaction of the two sRNAs? We are currently carrying out experiments to study the steady-state levels of PhrS after OsiS induction, in the presence and absence of Hfq.

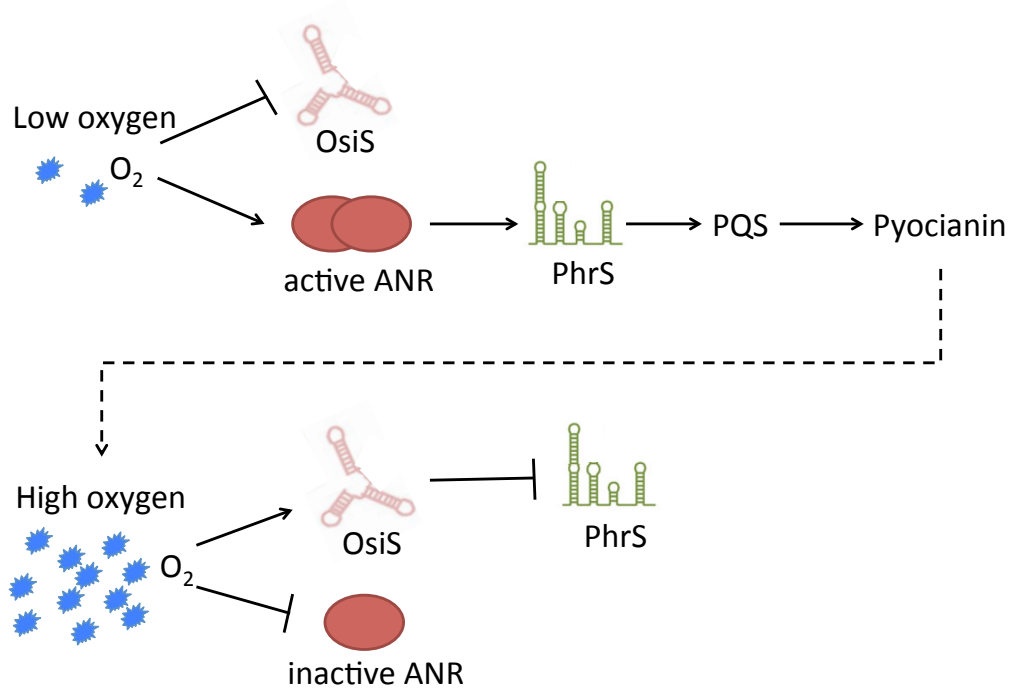


Figure 5. Current model of PhrS regulation by the sRNA OsiS. When oxygen is limited, there is no expression of OsiS and the ANR protein gets activated. The activated ANR is required for the transcription of the *phrS* gene, which leads to a higher production of PQS. PQS activates the production of pyocianin, which can react with molecular oxygen to generate oxidative stress. Under conditions that generate oxidative stress, OsiS is highly expressed, which produces a severe downregulation in the levels of PhrS. In addition, ANR is not active and does not promote the synthesis of PhrS.

In addition to its function as a regulatory RNA, PhrS encodes a highly conserved 37 amino-acid peptide (7). The peptide contains a predicted transmembrane segment and sub-cellular fractionation revealed that the peptide is indeed located in the cytoplasmic membrane (7). However, the PhrS peptide does not affect the release of PQS, and it is therefore unknown whether it acts within the same regulatory circuit as the PhrS sRNA or even if it has any cellular function (13). OsiS could be downregulating PhrS in order to lower the levels of the PhrS peptide, and not only affect the PhrS base-pairing function. Moreover, OsiS could be regulating some protein or some other transcripts at the level of translation and therefore its effect could not be assessed by RNA-Seq. Proteomic analysis are currently being carried out to address these questions.

Under hypoxic conditions, PQS production is limited. PhrS is assumed to counteract the low levels of PQS at reduced or fluctuating oxygen tensions by stimulating the expression of *pqsR* (13). Thus, the role of sRNA OsiS is to "counteract the counteracting effect" of sRNA PhrS at high oxygen tension. Thus both OsiS and PhrS respond to oxygen levels and allow the fine-tuning of PQS synthesis. So what is the biological implication of oxygen-dependent fine-tuning of PQS? In laboratory conditions, the growth rate of mutants deficient in PhrS production is the same as the wild-type strain, though the deletion mutant produced less

pyocianin (13). We could not produce a strain lacking the OsiS gene, however over-expression of OsiS does not affect any apparent phenotype. During infection, *P. aeruginosa* is subject to constant shifts in oxygen tension due to the human airway structure, biofilm formation, antibiotic treatment and the host immune response. *P. aeruginosa* faces oxygen limitation when growing in the lower airway mucus of CF patients (51). In addition *P. aeruginosa* can form biofilms that contain multicellular structures that show strong oxygen gradients (52-54). Some antibiotics routinely used against *P. aeruginosa* infections, like ciprofloxacin, produce an increase of ROS in the bacterial cells (25,26). A variety of studies have shown that *P. aeruginosa* encounters ROS in the lungs of CF patients due to the exaggerated, sustained and extended inflammatory response, characterized by influx of neutrophils and high concentrations of interleukin-8 (27-29,55-57). We infer that *P. aeruginosa* most likely needs a tight and fine-tuned regulation of the mechanisms that allow adapting to shifts in oxygen tension, and that the sRNA OsiS is involved in this kind of regulation by responding to high oxygen tension and regulating the levels of sRNA PhrS. Thus, fine-tuning regulators as PhrS and OsiS make no extraordinary difference in controlled and steady laboratory conditions, however we hypothesized that they might be pivotal for a fast and fine-tuned regulation of the mechanisms necessary to survive during adaptation to an environment with such fluctuating levels of oxygen tension as the one inside the host.

ACKNOWLEDGEMENTS

This work was supported by the Novo Nordisk Foundation and the Danish National Research Agency. Søren D. Pedersen is gratefully thanked for assistance in the design of the cloning strategy.

REFERENCES

1. Yang L, Jelsbak L, Marvig RL, Damkiær S, Workman CT, Rau MH, et al. Evolutionary dynamics of bacteria in a human host environment. Proc Natl Acad Sci USA. 2011 May 3;108(18):7481–6.

2. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, DArgenio DA, et al. Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. 2006 May 19;:1–6.

3. Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner E, Margalit H, et al. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Current Biology. 2001;11(12):941–50.

4. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, et al. Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq. PLoS Genet. 2008 Aug 22;4(8):e1000163.

5. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature. 2010 Mar 11;464(7286):250–5.

6. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. The small RNA

chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. Cell. 2004 Jul 9;118(1):69–82.

7. Sonnleitner E, Sorger-Domenigg T, Madej MJ, Findeiss S, Hackermüller J, Hüttenhofer A, et al. Detection of small RNAs in Pseudomonas aeruginosa by RNomics and structure-based bioinformatic tools. Microbiology (Reading, Engl.). 2008 Oct;154(Pt 10):3175–87.

8. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Genome-wide identification of novel small RNAs in Pseudomonas aeruginosa. Environmental Microbiology. 2012 Aug;14(8):2006–16.

9. Brencic A, McFarland KA, McManus HR. The GacS/GacA signal transduction system of Pseudomonas aeruginosaacts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. Molecular Microbiology. 2009 Aug;73(3):434–45.

10. Sonnleitner E, Abdou L, Haas D. Small RNA as global regulator of carbon catabolite repression in Pseudomonas aeruginosa. Proc Natl Acad Sci USA. 2009 Dec 22;106(51):21866–71.

11. Wilderman PJ, Sowa NA, FitzGerald DJ, FitzGerald PC, Gottesman S, Ochsner UA, et al. Identification of tandem duplicate regulatory small RNAs in Pseudomonas aeruginosa involved in iron homeostasis. Proc Natl Acad Sci USA. 2004 Jun 29;101(26):9792–7.

12. Oglesby-Sherrouse A, Vasil M. Characterization of a heme-regulated non-coding RNA encoded by the prrF locus of Pseudomonas aeruginosa. PLoS ONE. 2010;5(4):e9930.

13. Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, et al. The small RNA PhrS stimulates synthesis of the Pseudomonas aeruginosa quinolone signal. Molecular Microbiology. 2011 May;80(4):868–85.

14. Gómez-Lozano M, Marvig RL, Molin S, Long KS. Identification of bacterial small RNAs by RNA sequencing. Methods in Pseudomonas aeruginosa. In press; 2013 Apr pp. 1–34.

15. Gonzalez N, Heeb S, Valverde C, Kay E, Reimmann C, Junier T, et al. Genome-wide search reveals a novel GacA-regulated small RNA in Pseudomonas species. BMC Genomics. 2008;9:167.

16. Livny J, Brencic A, Lory S, Waldor M. Identification of 17 Pseudomonas aeruginosa sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic Acids Research. 2006;34(12):3484.

17. Gallagher LA, McKnight SL, Kuznetsova MS, Pesci EC, Manoil C. Functions required for extracellular quinolone signaling by Pseudomonas aeruginosa. Journal of Bacteriology. 2002 Dec;184(23):6472–80.

18. DArgenio DA, Calfee MW, Rainey PB, Pesci EC. Autolysis and autoaggregation in Pseudomonas aeruginosa colony morphology mutants. Journal of Bacteriology. 2002 Dec;184(23):6481–9.

19. Price-Whelan A, Dietrich LEP, Newman DK. Rethinking “secondary” metabolism: physiological roles for phenazine antibiotics. Nat Chem Biol. 2006 Feb;2(2):71–8.

20. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology. 1995 Jul;177(14):4121–30.

21. Hoang TT, Kutchma AJ, Becher A, Schweizer HP. Integration-Proficient Plasmids for Pseudomonas aeruginosa: Site-Specific Integration and Use for Engineering of Reporter and Expression Strains. Plasmid. 2000 Jan;43(1):59–72.

22. Herrero M, de Lorenzo V, Timmis KN. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. Journal of Bacteriology. Am Soc Microbiol; 1990;172(11):6557–67.

23. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012 Apr;9(4):357–9.

24. Li H, Handsaker B, Wysoker A, Fennell T,

Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.

25. Becerra MC, Albesa I. Oxidative stress induced by ciprofloxacin in Staphylococcus aureus. Biochem Biophys Res Commun. 2002 Oct 4;297(4):1003–7.

26. Goswami M, Mangoli SH, Jawali N. Involvement of reactive oxygen species in the action of ciprofloxacin against Escherichia coli. Antimicrobial agents and chemotherapy. 2006 Mar;50(3):949–54.

27. Hull J, Vervaart P, Grimwood K, Phelan P. Pulmonary oxidative stress response in young children with cystic fibrosis. Thorax. 1997 Jun;52(6):557–60.

28. Doring G, Goldstein W, Botzenhart K, Kharazmi A, Schiøtz PO, Hoiby N, et al. Elastase from polymorphonuclear leucocytes: a regulatory enzyme in immune complex disease. Clin. Exp. Immunol. 1986 Jun;64(3):597–605.

29. Brown RK, Kelly FJ. Evidence for increased oxidative damage in patients with cystic fibrosis. Pediatr. Res. 1994 Oct;36(4):487–93.

30. Ding Y, Chan CY, Lawrence CE. Sfold web server for statistical folding and rational design of nucleic acids. Nucleic Acids Research. 2004 Jul 1;32(Web Server issue):W135–41.

31. Farmer JJ, Herman LG. Epidemiological fingerprinting of Pseudomonas aeruginosa by the production of and sensitivity of pyocin and bacteriophage. Appl Microbiol. 1969 Nov;18(5):760–5.

32. Blackwell CC, Winstanley FP, Telfer Brunton WA. Sensitivity of thermophilic campylobacters to R-type pyocines of Pseudomonas aeruginosa. J. Med. Microbiol. 1982 May;15(2):247–51.

33. Morse SA, Vaughan P, Johnson D, Iglewski BH. Inhibition of Neisseria gonorrhoeae by a bacteriocin from Pseudomonas aeruginosa. Antimicrobial agents and chemotherapy. 1976 Aug;10(2):354–62.

34. Campagnari AA, Karalus R, Apicella M, Melaugh W, Lesse AJ, Gibson BW. Use of pyocin to select a Haemophilus ducreyi variant defective in lipooligosaccharide biosynthesis. Infection and immunity. 1994 Jun;62(6):2379–86.

35. Jones LF, Thomas ET, Stinnett JD, Gilardi GL, Farmer JJ. Pyocin sensitivity of Pseudomonas species. Appl Microbiol. 1974 Jan;27(1):288–9.

36. Bakkal S, Robinson SM, Ordonez CL, Waltz DA, Riley MA. Role of bacteriocins in mediating interactions of bacterial isolates taken from cystic fibrosis patients. Microbiology. 2010 Jul;156(Pt 7):2058–67.

37. Ernst RK, D Argenio DA, Ichikawa JK, Bangera MG, Selgrade S, Burns JL, et al. Genome mosaicism is conserved but not unique in Pseudomonas aeruginosa isolates from the airways of young children with cystic fibrosis. Environmental Microbiology. 2003 Nov 27;5(12):1341–9.

38. Wolfgang MC, Kulasekara BR, Liang X, Boyd D, Wu K, Yang Q, et al. Conservation of genome content and virulence determinants among clinical and environmental isolates of Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences. 2003 Jul 8;100(14):8484–9.

39. Brazas MD, Hancock REW. Ciprofloxacin induction of a susceptibility determinant in Pseudomonas aeruginosa. Antimicrobial agents and chemotherapy. 2005 Aug;49(8):3222–7.

40. Chang W, Small DA, Toghrol F, Bentley WE. Microarray analysis of Pseudomonas aeruginosa reveals induction of pyocin genes in response to hydrogen peroxide. BMC Genomics. 2005;6:115.

41. Dorman CJ. DNA supercoiling and environmental regulation of gene expression in pathogenic bacteria. Infection and immunity. 1991 Mar;59(3):745–9.

42. Mavrodi DV, Bonsall RF, Delaney SM, Soule MJ, Phillips G, Thomashow LS. Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from Pseudomonas aeruginosa PAO1. Journal of Bacteriology. 2001 Nov;183(21):6454–65.

43. Déziel E, Lepine F, Milot S, He J,

Mindrinos MN, Tompkins RG, et al. Analysis of Pseudomonas aeruginosa 4-hydroxy-2-alkylquinolines (HAQs) reveals a role for 4-hydroxy-2-heptylquinoline in cell-to-cell communication. Proceedings of the National Academy of Sciences. 2004 Feb 3;101(5):1339–44.

44. McKnight SL, Iglewski BH, Pesci EC. The Pseudomonas quinolone signal regulates rhl quorum sensing in Pseudomonas aeruginosa. Journal of Bacteriology. 2000 May;182(10):2702–8.

45. Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics. 2008 Dec 15;24(24):2849–56.

46. Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. Cell. 1997 Jul 11;90(1):43–53.

47. Boysen A, Møller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. J. Biol. Chem. 2010 Apr 2;285(14):10690–702.

48. Durand S, Storz G. Reprogramming of anaerobic metabolism by the FnrS small RNA. Molecular Microbiology. Wiley Online Library; 2010;75(5):1215–31.

49. Chou JH, Greenberg JT, Demple B. Posttranscriptional repression of Escherichia coli OmpF protein in response to redox stress: positive control of the micF antisense RNA by the soxRS locus. Journal of Bacteriology. 1993 Feb;175(4):1026–31.

50. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The Listeria transcriptional landscape from saprophytism to virulence. Nature. 2009;459(7249):950–6.

51. Worlitzsch D, Tarran R, Ulrich M, Schwab U, Cekici A, Meyer KC, et al. Effects of reduced mucus oxygen concentration in airway Pseudomonas infections of cystic fibrosis patients. J. Clin. Invest. 2002 Feb;109(3):317–25.

52. Xu KD, Stewart PS, Xia F, Huang CT, McFeters GA. Spatial physiological heterogeneity in Pseudomonas aeruginosa biofilm is determined by oxygen availability. Applied and environmental microbiology. 1998 Oct;64(10):4035–9.

53. Yoon SS, Hennigan RF, Hilliard GM, Ochsner UA, Parvatiyar K, Kamani MC, et al. Pseudomonas aeruginosa anaerobic respiration in biofilms: relationships to cystic fibrosis pathogenesis. Dev. Cell. 2002 Oct;3(4):593–603.

54. Yang L, Nilsson M, Gjermansen M, Givskov M, Tolker-Nielsen T. Pyoverdine and PQS mediated subpopulation interactions involved in Pseudomonas aeruginosa biofilm formation. Molecular Microbiology. 2009 Dec;74(6):1380–92.

55. Konstan MW, Hilliard KA, Norvell TM, Berger M. Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. American Journal of Respiratory and Critical Care Medicine. 1994 Aug;150(2):448–54.

56. Khan TZ, Wagener JS, Bost T, Martinez J, Accurso FJ, Riches DWH. Early Pulmonary Inflammation in Infants with Cystic Fibrosis. American Journal of Respiratory and Critical Care Medicine. American Lung Association; 1995 Apr;151(4):1075–82.

57. Cantin A. Cystic fibrosis lung inflammation: early, sustained, and severe. American Journal of Respiratory and Critical Care Medicine. 1995 Apr;151(4):939–41.

Small RNA OsiS links oxidative stress to quorum sensing control in *Pseudomonas aeruginosa*

María Gómez-Lozano[1,2] and Søren Molin[1,2]*

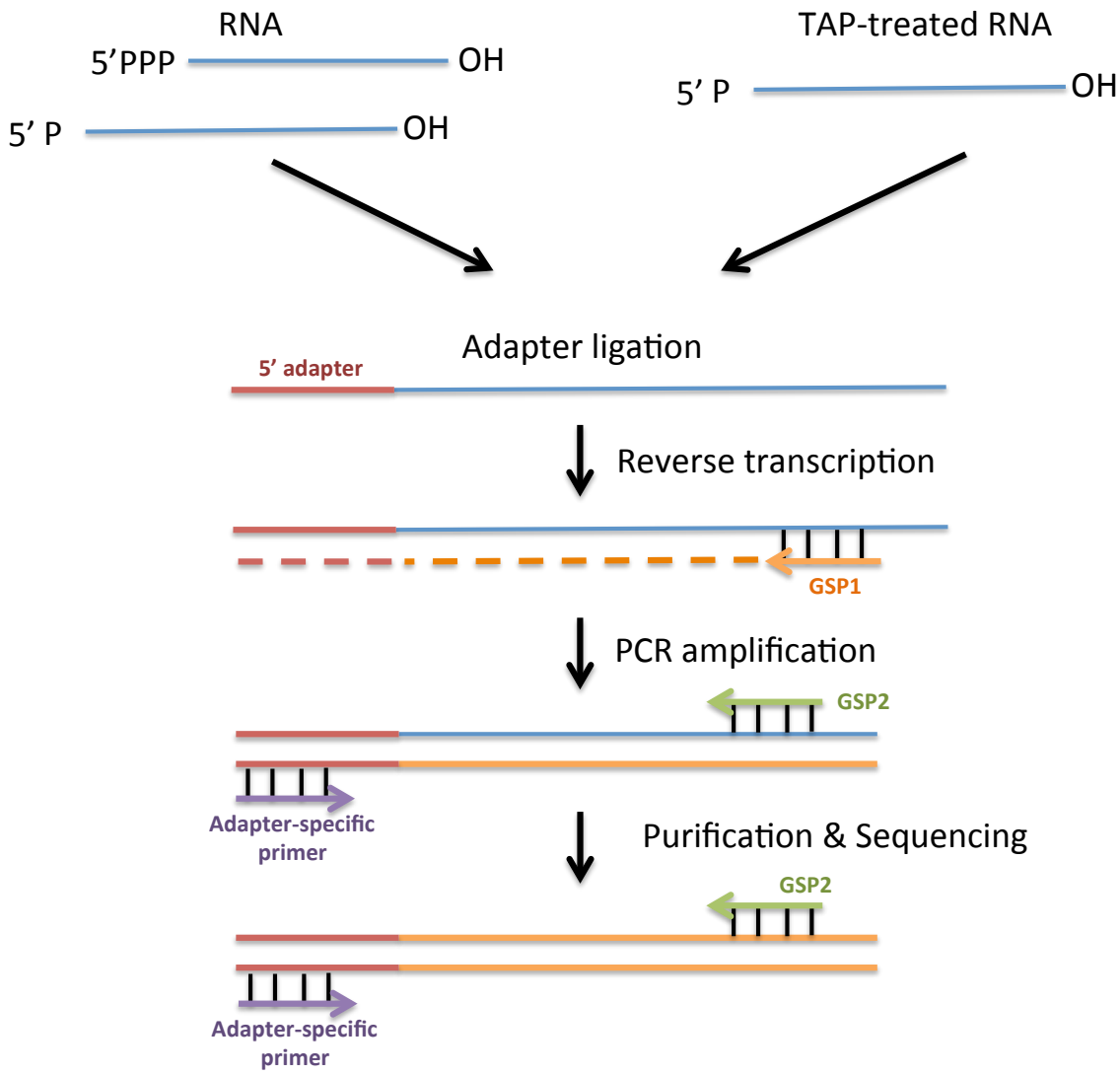
[1]*Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark.*

[2]*Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.*

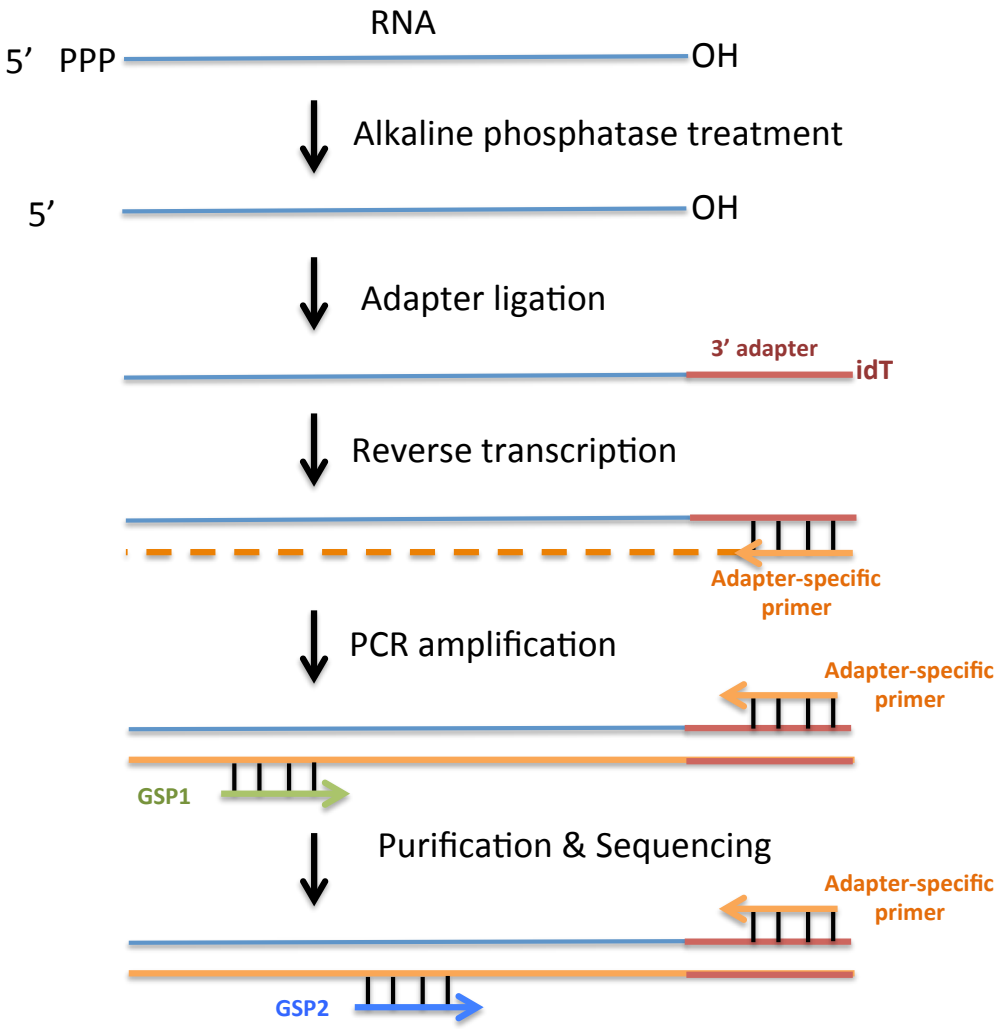
Supplementary material

Supplementary Table S1. Primers and adapters used in the study.

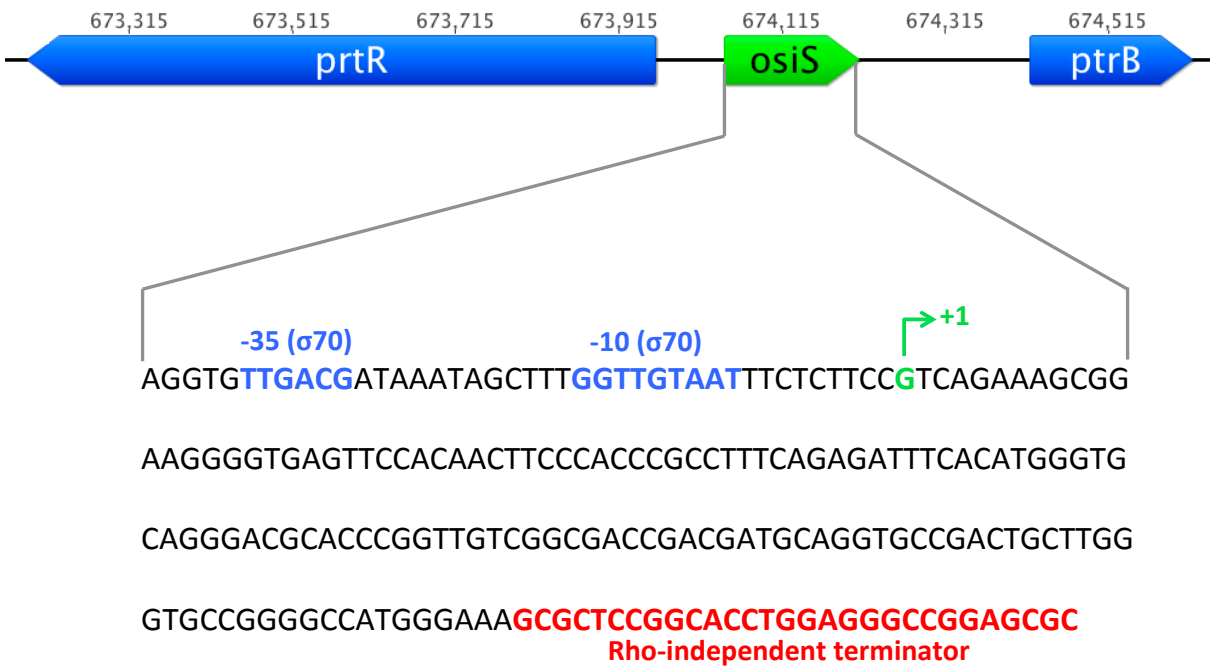
Primer	Sequence (5´->3´)	Use
23S-1954	AAAAAAAAAAAAAAAAAAACTTACCCGACAAGGAATTTCGC	Removal of 23S rRNA (provided in the MICROBExpress kit, Ambion)
23S-2511	AAAAAAAAAAAAAAAAAAGAGCCGACATCGAGGTGCCAAAC	Removal of 23S rRNA (provided in the MICROBExpress kit, Ambion)
16S-807	AAAAAAAAAAAAAAAAAATGGACTACCAGGGTATCTAATCC	Removal of 16S rRNA (provided in the MICROBExpress kit, Ambion)
16S-1114	AAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTACGGGACTT	Removal of 16S rRNA (provided in the MICROBExpress kit, Ambion)
5S	AAAAAAAAAAAAAAAAAAGCGTTTCACTTCTGAGTTCGGCA	Removal of 5S rRNA
F1-SacI	GAGAGAGCTCGGCCGCTTCTAGAGTTATGA	Plasmids pBAD-*osiS* and pBAD-control construction
R1_osiS	CACCCCTTCCGCTTTCTGACATGGAGAAACAGTAGAGAGT	Plasmid pBAD-*osiS* construction
F2_osiS	GTCAGAAAGCGGAAGGGGTGAGTTCCACAACTTCCCACCC	Plasmid pBAD-*osiS* construction
R2-XhoI_osiS	ACATCTCGAGTGGATGGAACATGACAGATG	Plasmid pBAD-*osiS* construction
R1_control	ATCAGGGCGTTCTTGATAAAATGGAGAAACAGTAGAGAGT	Plasmid pBAD-control construction
F2_control	TTTATCAAGAACGCCCTGATTCAGGACGGAGCGATAGATA	Plasmid pBAD-control construction
R2-XhoI_control	ACATCTCGAGTGGAATAGTGATGGCCACCA	Plasmid pBAD-control construction
Pser_up	CGAGTGGTTTAAGGCAACGGTCTTGA	Sequencing of pBAD-*osiS* and pBAD-control
Pser_down	AGTTCGGCCTGGTGGAGCAACTCG	Sequencing of pBAD-*osiS* and pBAD-control
5´ RNA adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA	5´-RACE
5´ adapter-specific primer	GCTGATGGCGATGAATGAACACTG	5´-RACE
5´-GSP1_osiS	AGGTGCCGGAGCGCTTTC	5´-RACE
5´-GSP2_osiS	CCCGGCACCCAAGCAGTC	5´-RACE
3´ RNA adapter E1	UUCACUGUUCUUAGCGGCCGCAUGCUC-idT	3´-RACE
3´ adapter E1 primer	CATGCGGCCGCTAAGAACAGTGA	3´-RACE
3´-GSP1_osiS	ACATGGGTGCAGGGACGCA	3´-RACE
3´-GSP2_osiS	GTTGTCGGCGACCGACGATGCA	3´-RACE
5S-F	ACGATCATAGAGCGTTGGAACCAC	RT-qPCR
5S-R	TGACGATGACCTACTCTCACATGG	RT-qPCR
osiS_F	AGGTGCCGGAGCGCTTTC	RT-qPCR
osiS_R	ACATGGGTGCAGGGACGCA	RT-qPCR



Supplementary Figure S1-A. 5´RACE strategy used to identify 5´ends of primary and processed transcripts.



Supplementary Figure S1-B. 3´RACE strategy used to identify 3´ends of transcripts.



Supplementary Figure S2. Organization of the *osiS* region in *P. aeruginosa* PAO1. The transcription start site of the primary transcript of OsiS was assessed by 5´-RACE.

Novo Nordisk Foundation Center for Biosustainability
Technical University of Denmark
Kogle Alle 6 2970 Hørsholm
Denmark

ISBN 978-87-91494-62-8

Technical University of Denmark